

Adhesion and Adhesives

K. W. Allen

*Joining Technology Research Centre,
Oxford Brookes University*

- I. History of Adhesion and Adhesives
- II. Mechanisms of Adhesion
- III. Classification of Adhesives
- IV. Solution-Based Adhesives
- V. Hot Melt Adhesives
- VI. Reactive Adhesives
- VII. Pressure-Sensitive Adhesives
- VIII. Surface Preparation for Bonding
- IX. Coupling Agents and Primers
- X. Joint Design
- XI. Failure of Joints
- XII. Test Methods

GLOSSARY

Adherend That to which an adhesive bond is being made; sometimes referred to as a substrate. (*Note:* Biologists use the latter term in a quite different sense.)

Adhesion An attraction between two substances which, after they have been brought together, requires work to be done to separate them. By convention, this excludes magnetic attractions.

Adhesive A substance which fills the gap between two surfaces and forms a relatively permanent and coherent bond (i.e., causes “adhesion” between them).

Curing An adhesive at an early stage in its use has to be a liquid so that it can wet and spread on the substrate, but then it has to become solid. This change from liquid to solid, however it is accomplished, is known as “curing.”

Phase A distinct form of matter: solid (crystalline, amorphous, etc.), liquid, or vapor.

Pre-treatment The preparation of an adherend surface to increase the strength and/or durability of adhesive bonding to it.

Solvent based Any solution could be described as “solvent based,” but in the present context the term is used for solutions in an organic solvent and implicitly excludes aqueous (water)-based solutions.

Surface tension/energy Because of the imbalance of forces at the surface of a liquid, the surface behaves as if it were a stretched skin. The tension in this skin, measured in force per unit length, is the surface tension. If the skin is stretched, then work has to be done and energy expended against that tension. This energy, in energy per unit area, is the surface energy. Provided that

a coherent system of units is used, the surface tension in dynes/cm or N/m is numerically identical to surface energy in erg/cm² or J/m².

Surfaces of high/low energy Generally, the surfaces of metals (or their oxides) are of high energy (200–1500), while most polymers are of low energy (20–50). Water is unusual, with a value of 72.

Vulcanize Originally used in rubber science to describe the cross-linking and eventual hardening from reaction with sulfur. Now extended to include any elastomer and its cross-linking or cure.

Wetting The extent to which a liquid will come into intimate contact with a surface to which it is applied. It may range from total spreading across the surface to remaining in droplets and making little or no contact. Good wetting of an adherend by an adhesive is a necessary, but not sufficient, condition for achieving a strong adhesive bond.

IT IS QUITE important to recognize the different meanings of the words in the title of this chapter before looking in detail at either. The topic of **adhesion** covers the scientific study of all the forces available to operate between two surfaces in intimate contact and which may prevent their separation without the application of some force. It encompasses a great many more phenomena than simply those of providing a practically useful joint between two components. The permanence of print on paper or plastic, the integrity of composites, the durability of paint, the way in which mussels or barnacles are attached to rocks or ship hulls, and the behavior of cancerous cells within the human body all involve adhesion (with no adhesives involved). In the reverse, the development of nonstick cookware and frictionless bearings requires an understanding of the basis of adhesion so that it may be avoided.

This science is of relatively recent development, with virtually no coherent work before the mid-1920s, although eventually some earlier concepts were found to be relevant in unexpected contexts.

In contrast, the technology of **adhesives** and their use dates back several millennia. Practical ways of uniting articles to provide a joint with resistance to an applied stress have been explored and developed to a useful level by entirely empirical methods. These are used in another wide range of instances, of which manufacturing plywood, securing cartons and packages, constructing aircraft, affixing postage stamps, and manufacturing carpets are all examples.

The technology of each of the applications is a mixture of (often very great) experience together with a modicum of understanding of some of the more elementary concepts from the science. Perhaps one of the more significant

changes over the recent past is the gradual swing of this balance toward a better understanding of why and how and away from tradition and rule of thumb.

I. HISTORY OF ADHESION AND ADHESIVES

The use of adhesion for practical purposes, either involving a separate adhesive or taking advantage of the phenomenon in other ways, goes back a very long way in human history (also, various examples may be found in both zoological and botanical structures). Probably the oldest reference to the use of an adhesive is in the Book of Genesis in the Bible, where we are told that the builders of the Tower of Babel used “slime” or “bitumen” as mortar. The earliest papyrus was made by using adhesion to bring the separate reed strands into a coherent film, and dates back to 3000 B.C. A little later, flour-and-water paste was used to join the separate sheets into a continuous roll. As the Egyptian civilization developed, from perhaps 1400 B.C., richly decorated artifacts were produced by using natural product adhesives to secure veneers and gems in place. While these techniques were used throughout the Imperial Roman Empire, they appear to have been lost with its fall (about 400 A.D.), and a dark period prevailed until the 16th century. At that time, adhesives began to be used for inlaying, veneering, and assembling furniture. However, even throughout the dark period, the technique of adhesion was preserved and used in various applications; for example, tar was used to make ships watertight.

In 1690, the first commercial plant to manufacture adhesives was opened in Holland, and this was followed in 1754 by the first patent being granted for making fish glue. Through the first half of the 19th century, development was slow, but gradually an increasing range of adhesives became available, including those using casein and natural rubber as base polymers. The development of synthetic resins really began with the production in 1909 of “Bakelite,” a phenol-formaldehyde resin, and later of urea-formaldehyde resin, both of which were used in the production of plywood. The importance of aircraft, accelerated by World War I, led to a considerable increase in the use of adhesives but no great innovations. These did not occur until the build-up to World War II and under the pressure of national crisis, when epoxy and urea/formaldehyde/formic acid (Aerolite, 1937) for wooden structures and aircraft and then the phenol-formaldehyde/polyvinyl formula (Redux, 1942) for aluminium aircraft adhesives were developed.

Since then, two successive forces have led to further developments. First has been the steadily growing

recognition of the advantages of using adhesive bonding in place of other methods of joining and fastening. These advantages include the absence of stress concentration, saving of weight, reduction of the need for surface refinishing (as there is no deformation from spot welding or rivets), and very quick curing for noncritical assembly. Consequently, there has been a growing demand for adhesives in new situations (e.g., automotive structures requiring enhanced toughness). To satisfy these demands, several new chemistries have been developed, such as anaerobic materials (1953), cyanoacrylates (1957), reactive acrylates (1976), and polyurethanes. These developments have been followed by an increasing awareness of the deleterious environmental effects of many of the materials used, particularly the organic solvents, which had been important constituents of many formulations. This has led to considerable efforts to produce improved, more environmentally friendly formulations, particularly water-based and hot-melt products.

II. MECHANISMS OF ADHESION

Explanations and theories of adhesion only began, in any serious way, relatively recently. In 1922, the Adhesives Research Committee of the British DSIR commented that there was "no generally acceptable explanation of the action by which glues cause surfaces to stick." In the past 80 years, however, numerous explanations and theories have been advanced. Initially, many of them suffered from the mistaken idea that a single explanation could be (and had to be) found to account for all examples of adhesion—whether sticking a postage stamp to an envelope or providing the strength necessary for an aircraft structure. Eventually the fallacy of this concept was realized, and a series of mechanisms have been recognized, each one having an appropriate part to play, either alone or in combination with others, in accounting for the huge range of examples of adhesion. These various mechanisms will be considered separately.

A. Mechanical Interlocking

A theory of mechanical interlocking is an intuitive concept and is the oldest explanation of adhesion. It is sometimes called "hooking" and involves the adhesive in liquid form flowing into various crevices and irregularities of the solid surface and hardening there. This gives a structure in which separation requires a mechanical force sufficient to break one of the components. Initially, the irregularities and crevices were considered to be of relatively large size, at least visible with an optical microscope. There are

instances where this notion seems reasonable—for example, classical dental fillings using amalgam and sticking of porous materials (e.g., textiles, paper, or cardboard). However, there are many instances where this does not seem possible (e.g., bonding of smooth, hard surfaces of glass or aluminium).

With the development of more sophisticated techniques for studying surfaces in the 1980s (ultra high-resolution electron microscopy), it became clear that these apparently smooth surfaces, particularly the aluminium used for aerospace structures, had both pores and protruding whiskers in the oxide layer which constituted their outermost surface. These structures were on a scale of Angstroms (10^{-10} m) rather than microns (10^{-6} m). The adhesive both penetrated into the pores and enclosed the whiskers to give a bond of great strength. Thus, the mechanical interlocking explanation received fresh support when a much smaller scale was considered.

B. Adsorption

The secondary molecular forces responsible for the phenomenon of physical adsorption of gases on solids also provide a source of adhesive strength. The forces are all the result of dipole interactions, Keesom dipole forces, Debye dipole/induced dipole forces, and London dispersion forces. These last arise from the instantaneous quantum mechanical asymmetry of electron clouds. They are universal and operate between every pair of materials. While they are relatively small, it can be shown that they are large enough to account for attractions far greater than any observed strengths. Thus, they are of universal significance, having some effect in every example of adhesive bonding, regardless of whatever other forces may or may not be involved. However, it is important to recognize that, like gases physisorbed on a solid surface, they are relatively easily disrupted. In particular, the adhesive is susceptible to displacement by water.

C. Chemical

There are a number of instances where, in addition to dispersion forces of interaction, there are chemical bonds involved. These may be either covalent bonds or have at least a considerable degree of ionic character. The methods for investigating these situations are sophisticated surface techniques (inelastic tunneling spectroscopy, XPS, and static SIMS), and they have been mainly employed in studying coupling agents and their properties. It is certainly true that some degree of chemical bonding enhances durability and resistance to environmental attack.

1. Acid-base interactions

Rather more than 20 years ago, it was recognized that in many situations there were both adsorption and other chemical bonding forces contributing to the ultimate bonding. Fowkes urged that all these—dispersion forces, various dipole interactions, and hydrogen bonding—should be summed to express the total interaction. He then went on to suggest that, with the exception of the dispersion forces, all the rest could be brought together and then considered as acid-base interactions following the Lewis acid-base scheme. This idea has been widely accepted, and considerable, more theoretical work dependent upon it has been done. However, there has also been criticism of the fundamental assumptions upon which this concept depends. The current (2001) situation is one that accepts this acid-base theory as being of significance and value, but it is not the ultimate, unique theoretical panacea in this discipline.

D. Diffusion

This theory is attributed to two Russian scientists: Voyutskii, who introduced the ideas and presented them in a significant monograph (1960–1963), and Vasenin, who developed the theory extensively. The concept is quite simple. If two surfaces of a high polymer above its glass transition temperature are brought together under a modest pressure to ensure intimate contact, then long chain molecules will inter-diffuse, and eventually the interface will be eliminated. This is now generally accepted for the adhesion of a rubbery material to itself, a process called “autohesion,” and is of importance in the production of automotive tires. However, attempts to extend the range of application of this explanation have been unsatisfactory and are now largely disregarded.

E. Electrostatic

This is a theory again due to a Russian scientist, Deryagin (1957 *et seq.*). He considered the case of a pressure-sensitive adhesive tape being peeled from a rigid substrate and regarded them as the two plates of a capacitance. An electric double layer forms and, in order to separate the two plates, work has to be done against the electrostatic attraction. While this hypothesis has been largely discounted, no other explanation has been advanced to account for the sparks or flashes of light which are certainly observed if an adhesive tape is rapidly stripped from a substrate under dry conditions.

F. Weak Boundary Layers

Although it is usual to discuss this concept of “weak boundary layers” under the heading of theories of adhe-

sion, it is really an explanation of the weakness of adhesive bonds rather than of their strength. It can easily be shown, on thermodynamic grounds, that a failure within either an adherend or the adhesive will always be energetically more favorable than a failure exactly at the interface. Because, not infrequently, failures appear to occur exactly between the adherend and the adhesive layer, Bikerman (1961) suggested that there was an imperceptible weak boundary layer at the surface and it was this weak layer which was the site of the failure. He argued that this was the site and cause of most failures. While this is not accepted as the predominant factor in joint failure, it is similar to the more recently developed concept of an “inter-phase” occurring between the two components of an adhesive joint.

III. CLASSIFICATION OF ADHESIVES

Adhesives may be classified in various ways: by chemical nature, by physical form, by relative strength of bond formed, or by mechanism of cure. For many purposes, this last is the most useful, since, with one notable exception, every adhesive has to pass through a phase transition in the course of its use. This change will involve one of three general processes:

1. Loss of solvent or dispersant
2. Solidification by cooling
3. Chemical reaction

In addition, there is one group that retains a liquid state, even if extremely viscous, throughout their useful life. Thus, adhesives are

1. Solvent based
2. Hot melt
3. Reactive
4. Pressure sensitive

Within each of these adhesives (but particularly the third, reactive), there are subdivisions corresponding to the particular chemistry involved.

While it is usual practice to describe an adhesive, particularly reactive adhesives, by the broad chemistry involved, it is important to recognize that all industrially satisfactory adhesives contain a range of additives to modify various properties (e.g., viscosity, stability, tack, gap filling). Thus, one very common one, which is apparently a simple solution of a single polymer, polychloroprene, in an organic solvent, actually contains a second polymer, a mixed solvent, a tackifying resin, a neutralizing agent, and an inert filler, at least; however, it is still usually referred to as a polychloroprene adhesive.

IV. SOLUTION-BASED ADHESIVES

The solution-based adhesives are either based on water and are, for the most part, strictly dispersions rather than true solutions, or are based on organic solvents, usually a mixture, and are true solutions.

One group of the water-based products has been in use for a very long time and involves natural products: carbohydrates from plants or protein materials from animal waste. The other group is based on synthetic materials developed largely to replace some of the solvent-based adhesives for environmental reasons.

A. Vegetable Adhesives

The largest group of vegetable adhesives is based on starch and its derivatives. They are still used in very large quantities worldwide, primarily in undemanding situations, because of their low cost. Unmodified starch, which is insoluble in cold water, has to be mechanically mixed with water and gently heated. The granules take up water reversibly, which is followed by a rapid swelling over a quite narrow temperature range. Finally, the water-filled sacs burst, and the dispersion now has maximum adhesiveness.

Dextrins result from a degradation of starch produced by heating, often with the addition of an acid. This has the effect of breaking 1–4 glucoside links and forming 1–6 links with consequent changes in molecular structure. These products provide a more easily used and reproducible adhesive than starch and are widely used in higher class paper bonding in the stationery trade.

There is also a small group of vegetable gums which have very particular and limited uses. These include gum arabic and gum tragacanth, which have some culinary and pharmaceutical uses.

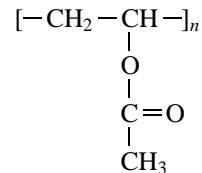
B. Animal and Related Protein Adhesives

The principal protein of bone, hide, skin, and sinew is collagen, which is insoluble in water but can be hydrolyzed in hot water. If only mild treatment of a relatively pure material is carried out, the product is gelatine, which is of relatively high molecular weight and is used for photographic films and in food. A more vigorous treatment of lower grades of raw materials yields a product that is darker in color and of lower molecular weight. This is concentrated to finally give solid cakes or powder. This is animal glue, which is used by dissolving in boiling water. Traditionally, this solution has been kept at about the boiling point in an inner container within an outer container of boiling water.

Slightly different grades are produced from different sources (bone or hide) but their properties and composi-

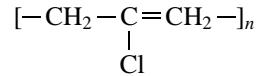
tions are all closely similar. In the manufacture of violins and similar instruments, it has been general practice to use only glue prepared from rabbit skins, but now some prepared from hare hide is being used. Although one prepared from fish bladder might be preferred, as it gives greater strength, apparently it is too brittle for general adoption.

C. Polyvinyl Acetate Adhesives



Probably the most important range of water-based adhesives are based on polyvinyl acetate. This is used as an emulsion which needs stabilizing. Most commonly, the principle stabilizer is about 5% polyvinyl alcohol. This is specially prepared by the hydrolysis of the acetate and retains about 20% of residual acetate groups in order to provide a molecular weight which is appropriately (relatively) low. This may be augmented by a secondary stabilizer, often gum arabic. This is widely used for woodworking and extensively for sticking paper and similar materials.

D. Polychloroprene Adhesives



One of the most widely used adhesives based on an organic solvent, both in manufacturing and in the do-it-yourself (DIY) field, is that based on polychloroprene. A simple solution of polychloroprene in toluene would produce a passable adhesive, but considerable improvements in bond strength and in resistance to creep can be attained by the addition of up to an equal amount of chlorinated rubber. This necessitates the addition of a proportion of an aliphatic ketone (acetone or methyl ethyl ketone) to the solvent. Furthermore, polychloroprene is always susceptible to "dehydrochlorination" with the liberation of free hydrochloric acid, and this is accelerated by the presence of chlorinated rubber. To combat this, a mixture of zinc and magnesium oxides is included. To improve tack a *tert*-butyl phenol/formaldehyde resin is included. This more complex mixture is used in the domestic DIY field and in many manufacturing plants, largely as a high-tack, "quick stick," contact adhesive. In the footwear industry, a similar product but without the tackifying resin is extensively used, being applied to both sides of the joint which are stored in that state. When the components are to be

assembled, the surfaces are reactivated by heating or by solvent wiping immediately before uniting them.

Generally, these adhesives quickly provide a bond of moderate strength which increases slowly as the polymer crystallizes. They have disadvantages in that the solvents are toxic, inflammable, and environmentally unacceptable. The footwear industry has now developed techniques which contain the solvent vapors and use them as fuel in the factory heating systems.

Considerable efforts are being made to replace these adhesives with aqueous-based products without unqualified success. The most promising techniques include the use of a foamed dispersion of the polymer in water and the use of a blend of two different polymers dispersed in water. A combination of polychloroprene and a polyurethane has given properties as good as, or in some cases better than, those obtained with polychloroprene alone.

V. HOT MELT ADHESIVES

Hot melt adhesives are one of the older types of adhesive. Sealing wax was used from medieval times. It was originally a mixture of beeswax, Venice turpentine, and a pigment, but later, when it was discovered, lac (or shellac) replaced the beeswax. Their use to any extent in industrial applications dates only from the 1950s for sealing food wrappings and milk cartons.

Simple waxes readily crystallized and so had very low cohesive strength. The first developments in their formulation were aimed at preventing crystallization and increasing their strength without interfering with the excellent wetting properties. To do so involved the inclusion of an elastomer as well as mixtures of synthetic waxes. With these improvements, there has been an enormous growth in their use, largely due to the freedom from adverse ecological effects. Initially, the industrial applications were confined to automatic machine processes. More recently, small-scale, hand-held guns have been developed and are now used in a huge range of applications.

A. Polyethylene and Ethylene-Vinyl Acetate Adhesives

An extensively used base material for hot-melt adhesives is polyethylene, because it is available in a wide range of molecular weights and because it is inexpensive. However, it requires additives to prevent crystallization as it solidifies. An equally popular base is the ethylene-vinyl acetate copolymers, which have the advantage of a degree of polarity which enhances its eventual adhesion and a considerably lower tendency to crystallize as it cools. These two polymers may be used, either alone or together, as the base combined with appropriate elastomer and antiox-

idant compounds. The application temperatures of these adhesives are commonly in the range of 160–190°C.

B. Warm Melt Adhesives

There are obvious advantages to be gained if the temperature of application of these adhesives could be significantly decreased without impairment of their adhesive properties. Considerable efforts have been made to this end with some measure of success. At least one manufacturer has produced a product based on a modified ethylene-vinyl acetate copolymer with an application temperature of 90–120°C with no loss of adhesive qualities.

C. Reactive Hot Melt Adhesives

Quite recently, a novel concept has been that of a hot-melt adhesive which could be caused to cross-link after it had been applied. This would give all the advantages inherent in hot melt products with the additional quality of a much higher service temperature. This has been achieved by the inclusion of terminal iso-cyanate groups in the structure of linear polyester molecules which are then applied as hot-melts. They then cross-link by the absorption and reaction of water with the iso-cyanate groups. Similar effects can be obtained by electron beam irradiation or by thermal decomposition (at a higher temperature than the application temperature) of a free radical initiator.

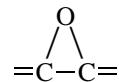
VI. REACTIVE ADHESIVES

In this group of adhesives, the change of phase from liquid to solid, which is an essential feature of all except "pressure-sensitive" adhesives, is brought about by a controlled chemical reaction. This reaction may follow any of several different mechanisms:

1. Reaction of two components of the adhesive system
2. Input of energy, which may be heat, electromagnetic radiation of various wavelengths, or electron beam
3. Reaction dependent upon some component of the environment, either its presence or its absence

A. Epoxides

These adhesives all depend upon linear compounds of moderate molecular weight which have pendent hydroxyl groups and are terminated with three-membered, oxygen-containing epoxy or oxirane rings.



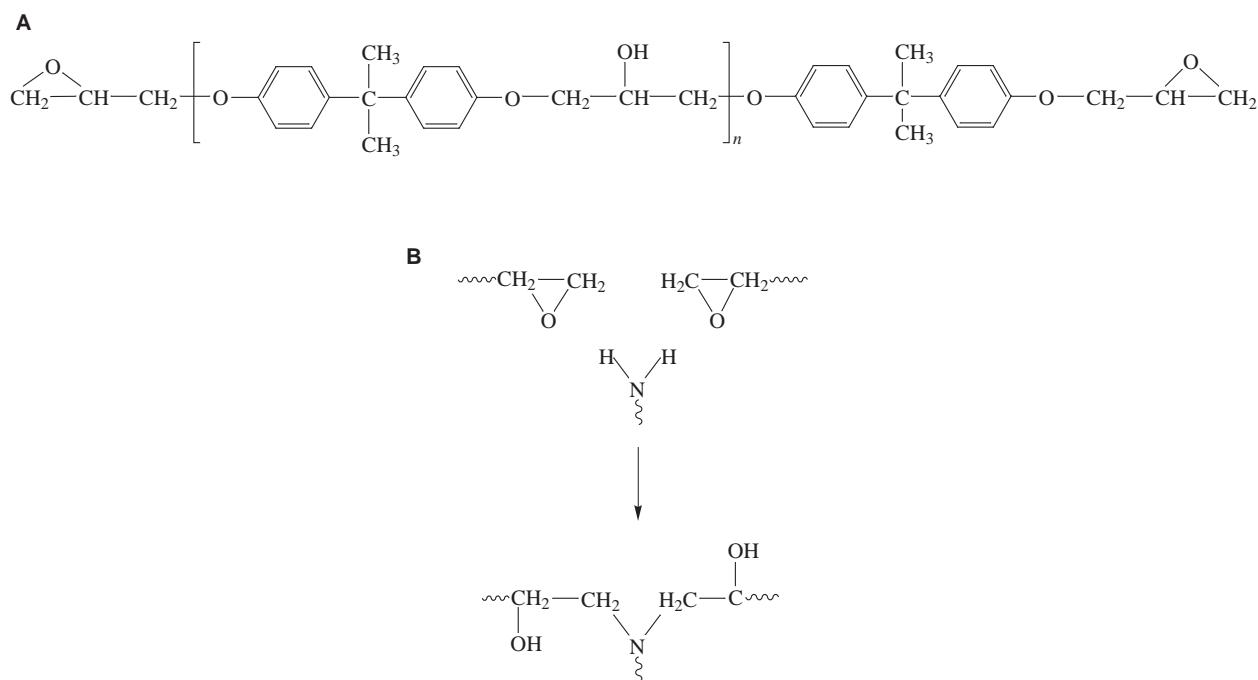


FIGURE 1 Epoxide adhesives and their cure. (A) The diglycidyl ether of bisphenol A (BADGE); the liquid resin grade suitable for adhesives has an average value of n of approximately 0.2. (B) Outline of cure with a primary amine; use of a diamine enables a fully cross-linked structure to be formed. Other curing agents (e.g., acid anhydrides) react to give similar results.

The polymerization or cure depends upon a reaction generally with either poly-functional amines or acid anhydrides. Thus, two components are, in one way or another, essential to their use. They react by addition polymerization, not condensation, so no volatile by-products are produced. The adhesive properties depend upon the hydroxyl groups; the epoxy groups are used in the cross-linking reactions.

While the epoxy component is usually a diglycidyl ether of bisphenol A (DGEBA, or commonly BADGE), there is a very wide range of second components available, apart from the various other ingredients (catalysts, fillers, toughening agents, etc.). The choice is dependent upon the particular requirements of the product.

Figure 1 outlines the main chemical reactions involved in the cure of epoxide adhesives. For practical purposes, the commercial products can be conveniently put into three groups:

1. Two-part, cold-curing formulations
2. One-part, heat-curing paste formulations
3. Film-form, hot-curing formulations

With the two-part, cold-curing adhesives, ranges of grades of the epoxy resin and of the chemistry of the curing agent are available. Generally, an additional catalyst is also needed to initiate the reaction. The choice of each

of the components for any particular product will depend upon the requirements: open time, service temperature, chemistry of the adherend, nature of the loading, etc.

A disadvantage of these formulations is that the proportions which have to be mixed immediately before application have to be controlled to within quite narrow limits. Thus, in any commercial process, fairly complex metering equipment is necessary.

These formulations are commonly treated to a post-curing heat treatment to increase their cross-linking and raise their ultimate strength. One-part, heat-cured paste formulations involve latent curing agents which are only released for reaction at elevated temperatures. Cure temperatures in the range of 120–180°C are recommended, and service temperatures of up to 180°C may be met. They will give higher joint strengths and retain useful joint strength at considerably higher temperatures than cold-curing formulations.

Film-form, hot-cured materials have been developed largely for the aerospace industry. In principle, these are similar to the one-part products but are produced in this form, often with a fabric support film.

All of these products may be toughened by the incorporation of very fine dispersed particles of an elastomer, very similar to the toughening of acrylic adhesives (*q.v.*). A comparison of characteristics of commercial epoxy adhesives is given in Table I.

TABLE I Characteristics of Commercial Epoxide Adhesives

Product	Cure temperature and time	Lap shear strength (MPa)	Service temperature
Two-part, cold-cured	Room-temperature	20	80–120°C
One-part, heat-cured	80–120°C, 30–60 min or more	31	<175
Film, heat-cured	120–170°C, >60 min.	38	<150

B. Acrylates

Several, largely unrelated series of adhesives are derived from acrylic acid, CH_2CHCOOH . These, with the approximate date of their introduction, include:

1. Anaerobic adhesives (1950)
2. Cyanoacrylates (“super-glues”) (1957)
3. Reactive acrylate adhesives (sometimes called “second-generation” acrylates) (1975)
4. Ultraviolet (UV) curing adhesives
5. Some miscellaneous adhesives, mainly solvent-based and of only minor significance

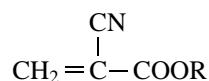
These will be considered in turn.

1. Anaerobics

The anaerobic adhesives are formulations in which polymerization is inhibited or prevented by the presence of oxygen. Once that has been excluded, then, particularly in the presence of ferrous ions, the liquid polymerizes to a hard solid. The base material is triethylene glycol dimethacrylate (TEGDMA), but the detailed chemistry is still somewhat obscure. As with virtually all adhesives, there are various additional components and catalysts that are necessary for a satisfactory product, and the production itself presents formidable problems. Much of the detail is to a considerable extent a commercial secret. The principal use of these products is for thread locking of threaded assemblies and co-axial joints where they act principally by mechanical jamming and only to a secondary extent as adhesives. Their use in these situations has resulted in very considerable savings of both time and money.

2. Cyanoacrylates

These compounds were introduced to the industrial scene in 1958 under the title “Eastman Kodak 910,” and it was not until the 1970s that their production and use expanded. They are all essentially esters of cyanoacrylic acid, thus:



where R is a lower aliphatic radical: methyl, ethyl, or, to a lesser extent, iso-propyl, *n*-butyl, or allyl.

The monomers are colorless, mobile liquids which very readily undergo anionic polymerization. This is initiated by relatively weak bases; the imperceptible film of adsorbed water on any metallic surface is quite sufficient. This is an extremely rapid reaction and is commonly completed within seconds.

Because of their ability to bond a wide variety of materials (even the user’s fingers!) very rapidly and easily, they are commonly known as “super-glues;” however, there are considerable limitations on their use. The monomers are very labile and may polymerize apparently spontaneously, so they have a very limited shelf life. Bonds are very readily disrupted by water and are only stable at temperatures less than 80°C. Nevertheless, they find extensive use in product assembly, especially in the electronic industry, where the final unit is likely to be exposed only to benign conditions. A somewhat unexpected application is in surgery, where it has some use as a skin suture. A great deal of effort has been expended in improving the toughness and stability of cyanoacrylates, with significant success, but they remain a product with limited uses and recognized limitations.

3. Reactive Acrylates

This group of adhesives depends upon two simultaneous but separate advances, the first of which is a graft polymerization reaction by a free radical path, which means that the proportion of the two components is much less important and does not need careful metering or control. The activator (or catalyst) can be surface applied and will react by diffusion into the monomer mixture. Also, a technique has been developed for greatly increasing the toughness of the cured product.

The adhesive comprises two parts. The first is a solution of an elastomer in a mixture of acrylate monomers together with an organic peroxide as a free radical generator. The elastomer is usually a chlorosulfonated polyethylene rubber (DuPont’s “Hypalon”) but alternatively may be a carboxy- or vinyl-terminated butadiene/acrylonitrile copolymer (B.F. Goodrich’s “Hycar”). This is a viscous gel and is commonly called the “adhesive.” The second part is a condensate of butyraldehyde and aniline with an accelerator and is a mobile liquid commonly called the “initiator.”

The adhesive is applied to one side of the joint and the initiator is applied to the other side; when the two are brought together they react quite quickly to give a

TABLE II Lab-Shear Joint Strengths of Various Structural Adhesives

	Lap-shear strength (MPa)
Conventional, two-part, cold-curing epoxy	24.1
	15.0
	13.2
PF/polyvinyl formal aircraft film adhesive cured at 165°C under pressure	28.9
Surface-initiated, toughened acrylics	31.0
	22.9
	19.2

strong bond. Commonly, handling strength is achieved in 1–3 minutes, full working strength in an hour, and ultimate strength in 24 hours.

A great advantage of these adhesives is that they have much greater toughness than other adhesives, so they can withstand very much greater impact loads. This has enabled them to be useful in various automotive structures as well as other situations demanding high performance. **Table II** shows typical test strengths of the major types of primary structural adhesives.

4. Ultraviolet Curing

A series of acrylate adhesives for special applications is cured by exposure to ultraviolet (UV) radiation. Obviously, they can only be used in situations where the adhesive can be exposed to the radiation—for example, if one or both adherends are transparent to these wavelengths. Little seems to have been revealed about the composition and mechanism of reaction of these products, although a great deal is known about photo-polymerization. However, three groups available are on the market:

1. Very simple mixtures of acrylic monomers
 2. Relatively sophisticated formulations of acrylic urethanes
 3. Highly sophisticated “rubber-toughened” acrylates

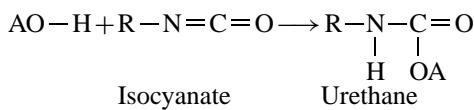
Products of this type, although not strictly adhesives, are extensively used in dentistry as fillings, often being built up in several layers.

While the chemistry of polymerization is different, there are parallels between these products and the anaerobic adhesives, and there are many similar problems in their formulation and production.

C. Urethanes

Unlike the majority of adhesives, the urethanes are not produced by a polymerization reaction but by the interac-

tion of an isocyanate with a hydroxyl-containing organic compound. The reaction is, in principle:

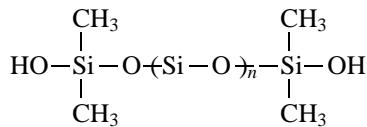


To produce useful products, both the isocyanate and the hydroxy compound are di- or poly-functional, so that cross-linking and chain extension can occur.

Because there are two radicals (A and R, above) which can be chosen independently, the properties of the polymer can be varied considerably (somewhat similar to the epoxides). The polymer produced can range from a stiff, bristle-like material to rubbers of various grades to a foam that is either flexible or rigid, all in addition to its use as a coating or an adhesive. It is also possible to incorporate compounds containing other active hydrogens, but the results are not urethanes; they are substituted ureas or products resulting from further reactions.

D. Silicones

Silicone (or siloxane) rubbers lie on the border between adhesives and sealants. They are based on linear poly(dimethyl siloxane) with silanol terminal groups:



where n is in the range of 300–1600.

These are combined with tri- or tetra-functional simple silane compounds as cross-linking agents. The common room-temperature vulcanizing (RTV) grades polymerize when moisture is absorbed from the surroundings and liberate either acetic acid or methanol, depending upon the radicals in the cross-linking agent. These are one of the most satisfactory sealants for a wide range of applications in many industries, including construction and automobile manufacture. While adhesive properties are necessary in these applications, the use of silicones primarily as adhesives is fairly limited.

E. Phenolics

This group of adhesives conveniently includes two somewhat different products:

1. Phenol-formaldehyde adhesives and their analogs, mainly used as wood adhesives
 2. Phenol-polyvinyl formal adhesives, mainly used in aircraft construction

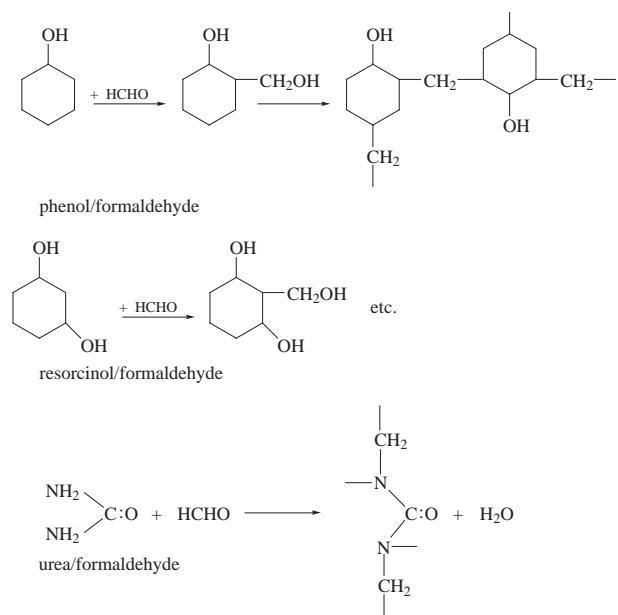


FIGURE 2 Synthesis of phenol/formaldehyde, resorcinol/formaldehyde, and urea/formaldehyde resins.

The resins derived from the interaction of phenol and formaldehyde were first produced in 1909 as a commercial product, “Bakelite,” a thermosetting, insoluble plastic; the reaction is outlined in Fig. 2. They were developed to produce adhesives, particularly for the manufacture of plywood, for which they were cured under pressure at elevated temperatures. For more demanding conditions (e.g., tropical or marine use), all or some of the phenol is replaced by resorcinol. Alongside this a somewhat similar series of polymers and adhesives based on the reaction of urea and formaldehyde was developed which had the advantage of lower cost but was less durable, so that its use is now largely confined to the production of particleboard for interior use.

The phenol-polyvinyl formal adhesives were developed during World War II in Britain and in parallel with similar poly-butyl products in the U.S. for metal aircraft construction. In the original formulation, a solution of phenol-formaldehyde resin was painted on a prepared metal surface and then powdered formal resin was sprinkled on. The whole structure was then transferred to an autoclave, where the adhesive cured under pressure (100 psi) with heat (150°C) for 30 minutes. In later developments, the product was formulated into a film form, which became the standard adhesive for aircraft construction and remains so, more than 50 years later. The long period of its use means that all its characteristics, both good and less desirable, are very well known and can be accommodated in design.

F. Polyimide and High-Temperature Adhesives

Epoxide adhesives are normally considered to have a upper useful temperature of about 100°C, although by careful formulation this may be increased to 150°C. There is a demand, particularly for military purposes, for adhesives that can withstand considerably higher temperatures, up to 300°C or even higher for limited periods. This demand has been met, at least to a limited extent, by some adhesive polymers based on combination of carbon and heterocyclic (nitrogen) ring systems. Of those that have been produced, the polyimides have achieved the greatest success, as they can retain more than 50% of their room temperature strengths at 300°C.

VII. PRESSURE-SENSITIVE ADHESIVES

The pressure-sensitive adhesives are different from all other types of adhesive. They are the familiar adhesive tapes used for packaging, masking, sealing, etc. (Sello-tape, Scotch tape), with a thin layer of adhesive spread on a polymer film backing. The fundamental difference is that they do not involve any change of phase. They begin as highly viscous, sticky liquids (viscosity of the order of 10⁶ poise) and remain so throughout their useful life. They must never cross-link or cure. The strength of their bond to a surface is dependent upon the pressure with which they are applied. To break the bond and peel off the tape requires the adhesive to flow and yield. If they cross-link and become hard and brittle, as can happen through exposure to sunlight in a window, the bond ceases to exist and the backing can be simply lifted away.

Originally, the adhesives used were based on natural rubber, and this is still a significant ingredient, but synthetic butyl rubber and poly(acrylate esters) are now gaining dominance. As with most adhesives, the optimum commercial product is a fairly complex mixture.

VIII. SURFACE PREPARATION FOR BONDING

A vital consideration in any use of adhesive bonding, before any bonding is attempted, is the state of the surfaces that are to be bonded. A considerable majority of the problems that arise with adhesively bonded joints can be traced to deficiencies in the surface preparation. It is fairly obvious that there must be no gross dirt or loose contamination; if there is, it must be removed. This can be easily accomplished by simple brushing, washing, or perhaps scrubbing. However, if bonds are to carry any significant load,

this is rarely sufficient. More persistent contamination may have to be removed and, depending upon the nature of the surface, the surface may require drastic modification—in some cases, the entire chemistry of the outermost layers has to be changed. Hence, it is necessary to review the various types of adherend surfaces individually.

A. Polymer Surfaces

Normally, the surface of polymers has little superficial contamination, and a simple wipe with an organic solvent will remove any traces. However, many of them are intrinsically difficult to bond because they are of low surface energy and are not wetted by any adhesive. This is particularly important for polyethylene (PE), polypropylene (PP), and polytetrafluoroethylene (PTFE), the various similar polymers containing fluorine. These all require drastic treatments which change the chemical nature of the surface. PE and PP may be oxidized by being passed through a gas flame, by corona irradiation, by plasma treatment, or by treatment with a chromic acid/sulfuric acid solution. Polytetrafluoroethylene requires defluorination, which is usually done by a chemical treatment but may be effected by various electric discharge techniques. The chemical method involves treatment with solutions either of metallic sodium in liquid ammonia or of sodium naphthalenide in tetrahydrofuran, both rather unpleasant reagents.

Polyvinyl chloride, having a higher surface energy, is less difficult; however, in its normal form, it contains a significant proportion of plasticizers, which migrate to the surface, but these can be removed with a solvent wipe.

B. Metal Surfaces

Because of their higher surface energies, these are far more prone to contamination with materials, particularly hydrocarbons, that are tenaciously held. The surfaces are covered with oxide films, many of which are only loosely attached to the substrate and are themselves of variable strength. In addition, their morphology is often not the most advantageous for strong and durable bonding. As a result, treatments specific to each metal have been developed.

1. Aluminum

Because of its use in aerospace equipment, treatment for this material has been developed more than for other materials. An outline of the recommended process is as follows. The first step is solvent degreasing, preferably by vapor condensation methods. This is followed by a sodium phosphate/detergent solution wash and then by etching in a chromic acid/sulfuric acid mixture. Finally, it is anodized in either a chromic acid or a phosphoric acid bath. Appropriate washes are applied between each step. At no stage

must the temperature exceed 62°C to avoid conversion of the aluminum oxide film from the desired bayerite form to boehmite

2. Steel

Frequently, the oxide (rust) on the surface of steel is friable and very weakly attached to the underlying surface. This oxide needs to be removed, usually by mechanical abrasion (e.g., wire brushing, grit blasting, abrasive papers). For many purposes, this is an adequate preparation for bonding. If not, however, then etching in a mixture of sulfuric acid (10% w/v) and oxalic acid (20% w/v) may be required. In this procedure, as in any etching process of steel, carbon is inevitably deposited and must be removed by “de-smutting,” either mechanically by wire brushing or by pickling in nitric acid.

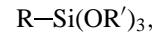
3. Titanium

Titanium and its alloys are particularly useful in aerospace structures because of their very advantageous strength-to-weight ratio and high service temperature; however, surface treatment for bonding has been less than satisfactory and has limited the use of titanium. A range of treatments has been devised over the past 40 years in an attempt to secure a bond with adequate durability, particularly under damp conditions. Chlorinated hydrocarbons are inadvisable for the initial degreasing because they may cause stress corrosion cracking, and in some circumstances a very violent reaction may occur. The recommended treatment is now a relatively simple sodium hydroxide anodizing at 10 volts in 20% w/v aqueous solution of NaOH at room temperature.

IX. COUPLING AGENTS AND PRIMERS

Coupling agents are used to improve the adhesion characteristics in bonds and in composite materials, especially in the presence of water. They were originally developed in the 1940s to improve the water resistance of glass fiber reinforced plastics (GFRP). Prior to their introduction, water would very rapidly penetrate along the interface between the matrix and fiber and destroy the integrity of the bond. Since their successful development, one important use of GFRP has been in the construction of small boat hulls.

The most significant group of compounds used in this way are the organo-silanes of the general formula



where R is a functional group which will interact with the adhesive or matrix; R' is usually methyl or ethyl, which, in use, is hydrolyzed to give a silan triol: Si(OH)_3 .

Alternative reagents, mainly used in the printing industry, are complex zirconates and titanates, but with

adhesives the silanols are almost exclusively used. They are used in two ways. The most common application is to apply them in a dilute aqueous solution to the substrate and allow them to dry before applying the adhesive and making the joint. Alternatively, a small proportion of the silanol may be incorporated in the adhesive, but this method is more commonly used in the coatings industry. This form of treatment is now used considerably to enhance the durability of adhesive joints in the presence of water. It is generally accepted that the mode of action is that the silane triol polymerizes to give a siloxane network covalently bound to the substrate surface.

While these materials are sometimes called "primers," that term is better reserved for various coatings which are applied to preserve surfaces that have been prepared for bonding but have to be stored before being used.

X. JOINT DESIGN

There are four ways in which an adhesive joint may be stressed:

1. Normally, to the bond line (either in tension or in compression)
2. In shear along the bond line
3. In cleavage at an angle to the bond line
4. By peeling one adherend away from the other

These are shown in Fig. 2.

Generally, adhesive bonds are strongest in compression and are of reasonable strength in tension or shear. But, while they can be satisfactory in tension, there is a considerable risk of joints intended to be in this mode having a considerable cleavage component through misalignment. Thus, this configuration in tension should be avoided. Cleavage and peel loadings are to be avoided whenever possible.

As a result of these factors, it is important that joints be specifically designed for adhesive bonding and that configurations originally intended for welding, riveting, or bolting are not used. Some illustrations of good and bad designs are given in Fig. 3.

XI. FAILURE OF JOINTS

A simple adhesive joint might appear to fail in five different sites: within the first adherend, at the interface between this adherend and the adhesive layer, within the adhesive layer, at the interface between the adhesive layer and the second adherend, or within the second adherend. Clearly, if the failure is within either of the adherends, the joint is more than sufficiently strong, but if it is within the adhesive then consideration must be given to whether an alternative adhesive would be better. A very considerable

proportion of failures, however, appear to be at one or other interface. Even apparently quite meticulous inspection may fail to reveal a thin layer of polymer on a metal surface, although more sophisticated techniques reveal its presence.

It can be shown theoretically that a parting exactly at an interface is energetically disadvantageous and hence improbable. Recent thinking has recognized that in the vast majority of cases there is an interphase between the two components which may be 100–300 molecules thick. It is somewhere within this interphase that failure occurs, so the failed surfaces may include material from both adherend and adhesive.

XII. TEST METHODS

Tests of adhesives and adhesive joints serve several purposes. These include:

1. Quality control of the adhesive and/or of the surface treatment
2. Determining basic mechanical and engineering properties of the cured adhesive as a material
3. Exploring the effects of particular environments on the joint
4. Selecting the most appropriate adhesive for a specific use
5. Enabling comparisons to be made between one manufacturer's product and another's

There are three commonly used test configurations: shear, tensile, and peel. Of these, the first, in the form of the single lap shear, is the most commonly used. It is easily made to precise specifications and quickly tested. It requires two metal pieces $4'' \times 1'' \times 1/16''$ wide, bonded with a $0.5''$ overlap. This joint is stressed to failure in the axis of the bond to give the failure load.

Tensile tests are of various configurations, but a common form is two cylindrical rods bonded together at right angles to their principle axis and stressed along that axis. The problem with this is the difficulty of ensuring perfect alignment, because even slight errors in this will impose cleavage stresses and invalidate the results.

A test that was devised to enable extensive testing of the effects of the environment is a wedge test, sometimes called the Boeing wedge test. In this test, two strips of metal are bonded together, face to face, and a wedge is driven into one end to force the bonded surfaces apart. The crack length which develops is measured as a function of time. Provided that the dimensions and mechanical properties of the adherends are known, it is possible to calculate the fracture energy.

When using any of these or other tests, it is very important to realize that the results are all dependent

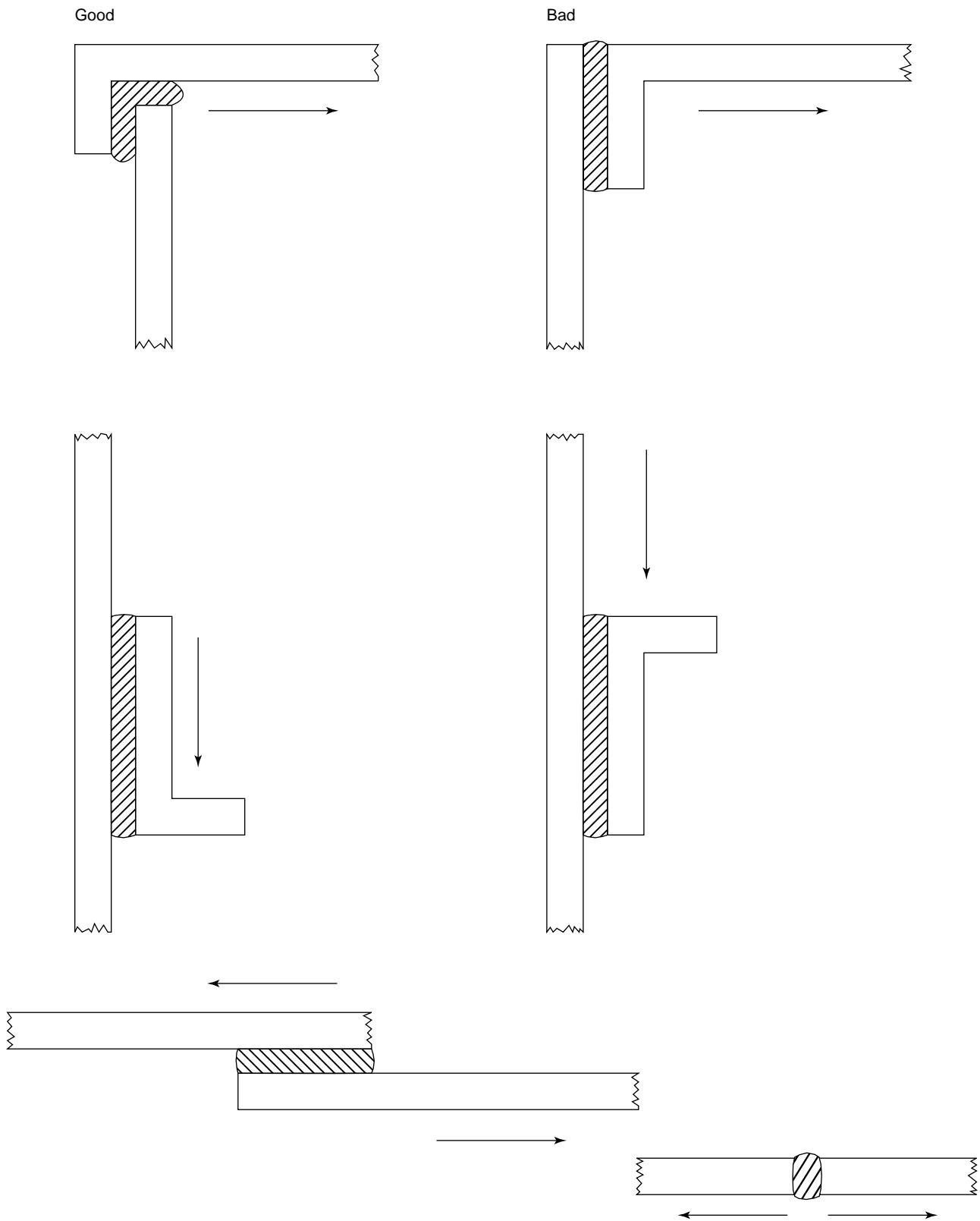


FIGURE 3

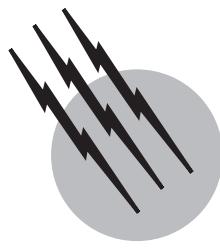
upon the entire system: the adhesive, surfaces of the adherends, the interface, and the interactions between all these. However, with care (and experience) the simple tests will give considerable insight into the questions that are posed.

SEE ALSO THE FOLLOWING ARTICLES

- ADSORPTION (CHEMICAL ENGINEERING)
- COATINGS, COLORANTS, AND PAINTS
- COMPOSITE MATERIALS
- ELECTROSTATIC POWDER COATING
- POLYMERS, MECHANICAL BEHAVIOR
- POLYMERS, THERMALLY STABLE
- SANDWICH COMPOSITES
- SURFACE CHEMISTRY

BIBLIOGRAPHY

- Adams, R. D., Comyn, J., and Wake, W. C. (1997) "Structural Adhesive Joints in Engineering," 2nd ed., Chapman & Hall, London, New York.
- Comyn, J. (1997). "Adhesion Science," Royal Society of Chemistry, Cambridge.
- Hussey, B., and Wilson, J. (1996). "Structural Adhesives—Directory and Databook," Chapman & Hall, London, New York.
- Kinloch, A. J. (1987). "Adhesion and Adhesives," Chapman & Hall, London, New York.
- Packham, D. E. (1992). "Handbook of Adhesion," Longman Scientific & Technical, Harlow/New York.
- Skeist, I. (1990). "Handbook of Adhesives," Van Nostrand-Reinhold, New York.
- Wake, W. C. (1982). "Adhesion and the Formulation of Adhesives," Applied Science Publishers, Barking.



Aluminum

Warren Haupin

Consultant, Lower Burrell, Pennsylvania

- I. History of Aluminum
- II. Natural Occurrence of Aluminum Compounds
- III. Mining of Bauxite
- IV. Extraction of Pure Alumina
- V. Electrolytic Reduction of Alumina to Aluminum
(Hall–Héroult Process)
- VI. Energy Considerations
- VII. Alternative Processes for Producing Aluminum
- VIII. Production of Ultra-High-Purity Aluminum
- IX. Properties of Aluminum
- X. Alloys of Aluminum
- XI. Forming and Shaping Operations
- XII. Finishes for Aluminum
- XIII. Uses of Aluminum
- XIV. Environmental Considerations

GLOSSARY

- Anode effect** Condition in which the electrolyte no longer wets the anode, creating a high electrical resistance.
- Bauxite** Rock or earthy soil that contains a high concentration of aluminum hydroxide minerals.
- Bayer liquor** Recycled aqueous solution of sodium hydroxide (NaOH) and sodium aluminate (NaAlO_2).
- Bayer process** Process to extract pure alumina (Al_2O_3) from bauxite.
- Cold work** Shaping metal by pressure at room temperature.
- Desilication** Removal of soluble silicates from Bayer

liquor by forming insoluble sodium aluminum silicates called DSP (desilication product).

Dewet To change a wetted surface to a nonwetted surface.

Digestion High-temperature dissolution or extraction.

Dolime Calcined dolomite rock ($\text{CaO} + \text{MgO}$).

Grains (in metal) Microscopically resolvable crystals.

Guinier-Preston (GP) zones Clusters of solute atoms in a supersaturated solid-solution lattice.

Hall–Héroult process Process for electrolytic reduction of alumina dissolved in molten cryolite.

Ingot Cast block of metal suitable for mechanical shaping by rolling, forging, extruding, and similar processes or for remelting.

Magnetohydrodynamic stability Stability of a fluid carrying an electric current in a magnetic field.

Overboltage Additional voltage over the equilibrium value for an electrode reaction to proceed at a finite rate.

Pig Cast blocks of metal designed for remelting.

Pot Smelting plant jargon for electrolytic cell.

Primary aluminum Virgin or new aluminum.

Reduction 1. Extraction of a metal from its ore. 2. Making ingot, plate, sheet, or rod, thinner by rolling, forging, drawing (stretching), and so on.

Scalping Smoothing the surface of an ingot by cutting away the surface imperfections.

Secondary aluminum Recycled aluminum scrap.

Smelting Chemical or electrochemical reduction of an ore; also called winning.

Wrought Product that has been shaped by mechanical working, as by rolling, extruding, or forging.

ALUMINUM is the most abundant metallic element, making up more than 8% of the earth's crust. Its lightness, high thermal and electrical conductivity, corrosion resistance, workability, and strength when alloyed make it the most widely used nonferrous metal. Large quantities of aluminum are used in building and construction, transportation (land, air, and sea), beverage cans, other packaging applications, electric power transmission lines, consumer durables, equipment, and machinery. The aluminum industry is second only to the steel industry in volume of metal produced and is the world's largest electrochemical industry. The principal ore, bauxite, occurs predominantly in the equatorial regions, where it was formed by weathering of igneous rocks. This produced a concentrated ore, containing 20–30% aluminum as oxide and hydroxides. Pure (>99%) alumina (aluminum oxide) is extracted from bauxite by the Bayer process. This involves dissolution in a hot sodium hydroxide solution, which leaves behind the impurities, mainly hematite (Fe_2O_3), titania (TiO_2), and silica (SiO_2), as an insoluble residue. Pure alumina is then precipitated from the solution as a hydroxide, filtered, washed, and then calcined to pure alumina (Al_2O_3). Aluminum metal is extracted from the alumina electrolytically by the Hall–Héroult process. In this process, the purified alumina is dissolved in an electrolyte consisting mainly of molten cryolite. Consumable carbon anodes are employed, producing carbon dioxide and carbon monoxide, which escape from the cell while the molten aluminum accumulates at the cathodic bottom and is siphoned out periodically. This aluminum is then alloyed to increase its strength and either cast directly into products or cast into ingots, which can then be formed into sheet or bar stock by rolling or into various products by forging, drawing, or extruding. Aluminum can be machined readily by conventional metalworking equipment. Sheet

can be stamped, cold-formed, bent, deep-drawn, stretched, ironed, embossed, and swaged. Aluminum can be joined by welding, brazing, soldering, and adhesive bonding and by mechanical means such as crimping, seaming, riveting, bolting, and screwing. It accepts a wide variety of finishing techniques, including painting, plating, anodizing, porcelain enameling, etching, chemical conversion, and mechanical texturing. Some aluminum is atomized to produce aluminum powder for pigments, powder metallurgy, explosives, or rocket fuel.

I. HISTORY OF ALUMINUM

In 1746 J. H. Pott prepared pure alumina from alum, long known as a naturally occurring compound. The oxide was called alumina because it was derived from alum. As chemists began to recognize that alumina was the oxide of a metal that had never been seen, a concerted effort was made to isolate the metal. Hans Christian Oersted probably was the first to produce metallic aluminum. In 1824 he prepared anhydrous aluminum chloride by passing chlorine through a heated mixture of charcoal and alumina. The aluminum chloride vapor condensed in a cooler region of his system protected from air. He then reacted a potassium amalgam with the aluminum chloride to produce an aluminum amalgam. By distilling off the mercury under vacuum, he obtained a small quantity of a metal that he reported "in color and luster resembled tin." Unfortunately, he published his work in an obscure Danish journal, and as a result credit for the discovery of aluminum is generally given to Friedrich Wöhler. In 1827 Wöhler reacted metallic potassium with anhydrous aluminum chloride in a porcelain crucible and obtained aluminum. The metal, however, remained a laboratory curiosity until 1854, when Saint-Claire Deville demonstrated that less costly metallic sodium could be substituted for potassium as the reductant and that the sodium chloride formed by the reaction produced a low-melting NaAlCl_4 complex that acted as a flux to coalesce the aluminum droplets. This led to the commercialization of the process in 1855. Deville added calcium fluoride and later cryolite (Na_3AlF_6) to the melts and was first to record that fused cryolite served as a solvent for the aluminum oxide that formed on the aluminum. Improved lowercost methods for producing sodium and sodium aluminum chloride were developed by Castner of New York. By 1889, 500 lb of aluminum per day, selling for 16 shillings per pound, were being produced by the Deville–Castner process at a plant near Birmingham, England.

In 1884, the pyramidal cap for the Washington Monument was cast from aluminum produced by sodiothermic reduction of NaAlCl_4 by Colonel William Frishmuth in Philadelphia. The 2.8 kg cap is still in place and in good condition after a century of exposure to the atmosphere.

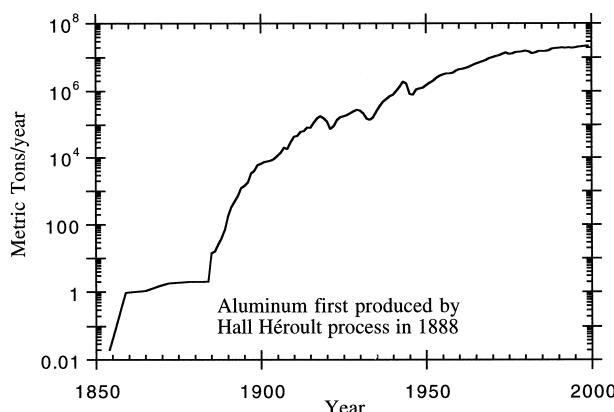


FIGURE 1 Annual world production of primary aluminum, 1850–2000 [Drawn from data in Grjotheim, K., et al. (1982). “Aluminium Electrolysis,” p. 6, Aluminium-Verlag GmbH, Dusseldorf. Extended data from U.S. Bureau of Mines, Mineral Commodity summaries, 1982–1999].

Deville in France and also Bunsen in Germany produced small quantities of aluminum electrolytically in 1854 by electrolyzing a sodium chloride–aluminum chloride melt in a porcelain container using carbon electrodes. But the dynamo was not invented until later (1867); hence, these early attempts at electrolytic reduction were uneconomical.

In 1886 Charles Martin Hall in the United States and Paul L. T. Héroult in France discovered almost simultaneously and completely independently a process of electrolytic reduction of alumina dissolved in molten cryolite. Both of these scientists were guided by the work of Saint-Claire Deville. Hall received a patent on his invention in the United States and Héroult on his in France. In less than three years, both patents were being implemented. In November 1888, the first aluminum was produced by Hall at a plant in Pittsburgh, Pennsylvania. At about the same time, aluminum was produced by Héroult at Neuhausen, Switzerland. In August 1888, Karl J. Bayer, an Austrian chemist, was issued a German patent for an improved method of extracting alumina from bauxite. Also, at this time, low-cost electric power was becoming available from water power. Hall and Héroult’s patents, Bayer’s patent, and low-cost electric power stimulated the growth of the aluminum industry in Europe and America and quickly displaced the Deville–Castner process. The growth of the aluminum industry is shown graphically in Fig. 1.

II. NATURAL OCCURRENCE OF ALUMINUM COMPOUNDS

Aluminum is the third most abundant element in the earth’s crust but is never found free in nature. It is usually com-

bined with silicon and oxygen and frequently iron in rock or clay. Bauxite is a rich and abundant ore and became the principal source of alumina after Bayer’s discovery of an economical extraction process. Bauxite gets its name from Les Baux, a province in southern France where it was first found. The term *bauxite* is now applied generally to rocks that contain significant quantities of aluminum hydroxide minerals. It exists in many varieties having physical characteristics ranging from a dark brown ferruginous material to cream or light pink layers of hard crystalline gibbsite bauxite. Under favorable hydrological conditions, bauxite can originate from almost any alumina-containing rock. It is generally believed that most bauxites were formed by tropical weathering or laterization, a slow leaching process by which the more soluble compounds are carried away by circulating ground waters in alternate dry and wet seasons. High temperatures, a pH between 4 and 10, heavy vegetation, bacteriological action, high permeability, a mature topography permitting accumulation of weathered products, and free movement of the water table with minimum erosion intensify laterization. Bauxites generally consist of mixtures of the minerals gibbsite [$\text{Al}(\text{OH})_3$], boehmite, and diasporite [$\text{AlO}(\text{OH})$]; clay minerals such as kaolinite [$\text{Al}_2\text{Si}_2\text{O}_5(\text{OH})_4$]; quartz (SiO_2); and anatase, rutile, and brookite (TiO_2). Small amounts of magnetite (Fe_3O_4), ilmenite (FeTiO_3), and corundum (Al_2O_3) are sometimes present. In addition to these minerals, bauxites usually contain traces of oxides of zirconium, vanadium, gallium, chromium, and manganese. The term *alumina trihydrate* is often applied to the mineral gibbsite, and *alumina monohydrate* to the minerals boehmite and diasporite. These terms are misnomers because the minerals are not true hydrates. There are bauxite deposits of gibbsite from the Tertiary or more recent periods. In older bauxites, such as the Mesozoic deposits in the Mediterranean region, gibbsite is dehydrated and altered to boehmite, which often is altered further to diasporite in the oldest known deposits. Factors other than age also affect composition, and many deposits contain combinations of the various types. The suitability of a bauxite for alumina extraction by the Bayer process depends not only on its alumina content but also on its content of combined silica, usually as kaolinite. Kaolinite not only contains aluminum that cannot be extracted, but also reacts with the caustic aluminate solution, producing a loss of sodium hydroxide. Worldwide bauxite reserves are given in Table I.

While bauxite is the richest ore, there are many alternative sources of alumina (Table II). Although the amounts are huge, the lower concentration of aluminum compounds and greater difficulty of processing them make these alternative materials less desirable. Coal refuse, particularly fly ash, presents an interesting possibility. The

TABLE I Bauxite Production (Thousand Metric Tons, Rounded) and Reserves (Million Metric Tons)

Country	1980	1985	1990	1995	1998	Reserves
Australia	27,179	32,400	41,400	42,700	45,000	3,200,000
Brazil	4,152	6,433	9,700	10,200	12,500	3,900,000
China	4,700	2,000	2,400	5,000	8,500	720,000
Guinca	13,911	14,738	15,800	15,800	16,500	7,400,000
Guyana	3,052	2,153	1,400	2,000	2,600	7,00,000
India	1,785	2,127	4,900	5,200	6,000	1,500,000
Jamaica	12,064	6,219	10,900	10,900	12,606	2,000,000
Russia	6,400	6,200	9,300	3,100	3,400	200,000
Suriname	4,903	3,375	3,500	3,500	4,000	580,000
United States	1,559	674	0	0	0	20,000
Venezuela	0	0	800	5,000	4,500	320,000
Other Countries	16,101	13,244	12,900	8,600	9,370	4,100,000
Total	92,806	89,557	113,000	112,000	125,000	25,000,000

American Bureau of Metal Statistics, Inc. Secaucus, NJ 1980, 1985 U.S. Geological Survey, Mineral Commodity Summaries, 1990–1998.

cost of mining is already paid and the ash presents a disposal problem, but the high temperature to which it has been heated makes extracting the alumina difficult.

III. MINING OF BAUXITE

In the mining of bauxite, first the size and shape of the deposite are determined by core drilling. Trees, if present, are harvested and the overburden removed and set aside. Then the bauxite is broken up, often with the aid of explosives, and removed with conventional earthmoving equipment. It is then crushed and washed to remove clay and silica, where this operation is economical, dried, and shipped to the Bayer process refinery. After the bauxite has been removed, the land must be restored to environmentally acceptable grade and contour and the overburden restored.

TABLE II Alternative Sources of Alumina in the United States^a

Material	Resource (megatons)
Low-grade Bauxite	36
Kaolin	320
Other clays	760
Shale	64
Alunite	32
Dawsonite	3,300
Anorthosite	15,000
Coal refuse	3,300

^a From *Bull. U.S. Bur. Mines* 667 (1975) and IC8335 (1967) and *U.S. Geol. Surv. Pap.* 820 (1973).

This land is often vulnerable to erosion, so careful planting is necessary to preserve the topsoil.

IV. EXTRACTION OF PURE ALUMINA

The world production of alumina was more than 37 million tons in 1988 (Table III). All of it was produced by the Bayer process and practically all was processed from bauxite. Technology has been developed to extract alumina from clays, anorthosite, alunite, nepheline, leucite, and other minerals. However, producing alumina from these other minerals is more costly and more energy intensive than processing bauxite.

In the Bayer process (Fig. 2), bauxite received from the mines is crushed to a particle size of less than 25 mm, usually by a hammer mill, blended to obtain a more uniform feed for the process, and stored. From storage the bauxite is metered into wet grinding-rod mills, where spent liquor (recycled NaOH solution) is used as a grinding medium. Lime (CaO) or sometimes do-lime (CaO + MgO) is added to assist in the extraction of alumina, to scavenge impurities, and later to enhance clarification. The slurry then flows to agitated storage tanks, where it remains for 20–30 hr to allow preliminary desilication to take place by the formation of insoluble sodium aluminum silicate. The slurry is then metered into high-temperature digesters, where alumina is extracted from the bauxite as sodium aluminate (NaAlO_2):



TABLE III World Production of Alumina (Thousand Metric Tons)

Country	1980	1985	1990	1995	1997
Australia	7,246	8,789	11,200	13,147	13,385
Brazil	493	882	1,660	2,141	2,756
Canada	1,202	1,019	1,090	1,064	1,165
China	750	800	1,460	2,200	3,000
France	1,173	877	606	425	350
Germany	1,651	1,741	972	825	750
Greece	494	487	587	598	602
Guinea	708	563	631	616	640
Hungary	231	801	826	184	76
India	805	569	1,600	1,650	1,700
Ireland	500	650	885	1,186	1,200
Italy	900	555	752	857	850
Jamacia	2,456	1,512	2,870	3,030	3,411
Japan	1,936	1,409	481	363	340
Kazakstan	<i>b</i>	<i>b</i>	<i>b</i>	1,022	1,050
Romania	534	480	440	323	282
Russia ^a	2,700	4,200	5,900	2,300	2,300
Spain	58	742	1,000	1,070	1,110
Surinam	1,316	1,369	1,530	1,589	1,600
Turkey	138	125	177	172	159
Ukraine	<i>b</i>	<i>b</i>	<i>b</i>	1,100	1,000
United Kingdom	102	106	115	108	100
United States	6,810	4,717	5,230	4,533	5,090
Venezuela	0	1,120	1,290	1,641	1,800
Yugoslavia	1,058	1,135	1,090	<i>b</i>	<i>b</i>
Other Countries	228	—	175	226	375
Total	33,489	34,648	42,600	42,370	45,085

^a USSR before 1992.^b Not applicable.

From U.S. Geological Survey.

Trihydrate bauxites can be processed in 30 min in liquor containing 120–135 g per liter Na₂O at ~140°C. The monohydrate bauxite require higher caustic concentrations (200–300 g per liter Na₂O), up to 4 hr of digestion, and temperatures of 200–300°C to obtain complete dissolution. Digestion is performed in steam-heated steel autoclaves or tubular reactors at pressures of up to 3.5 MPa. Further desilication takes place during digestion. At the end of digestion, pressure and temperature are reduced through a number of steps, flashing off steam. This steam is recycled for liquor heating. With the temperature now at ~105°C, the liquor loaded with undissolved particulate and dissolved NaAlO₂ passes to a sand clarifier, a conical bottom tank where the coarse, sandy residue settles out and is withdrawn continuously, washed, and transported to the bauxite residue disposal area. The overflow goes to thickeners, where starch and other flocculents are added to separate out fine mudlike particles. This mud is removed from the bottom of the tank as a 20–30% solid

slurry and sent to a countercurrent decantation bank and then either to a drying stack or to a bauxite residue lake. The clear, diluted, aluminum hydroxide-saturated liquor from the thickeners passes through filters and is cooled to between 47 and 62°C before precipitation. Aluminum hydroxide seed from the secondary and tertiary classifiers is added to the liquor as it enters the precipitation tank and the mixture agitated. The amount of seed may be up to four times the amount of aluminum hydroxide in solution. Reaction (3) takes place on the surface of the seed, causing the crystal to grow.



Precipitation is slow, requiring 20–80 hr of retention time. The liquor with precipitated Al(OH)₃ passes into three classification stages.

The first, or primary, classifier provides sufficient hydraulic lift that only the most coarse particles settle. These are removed as product slurried with bottom liquor. The

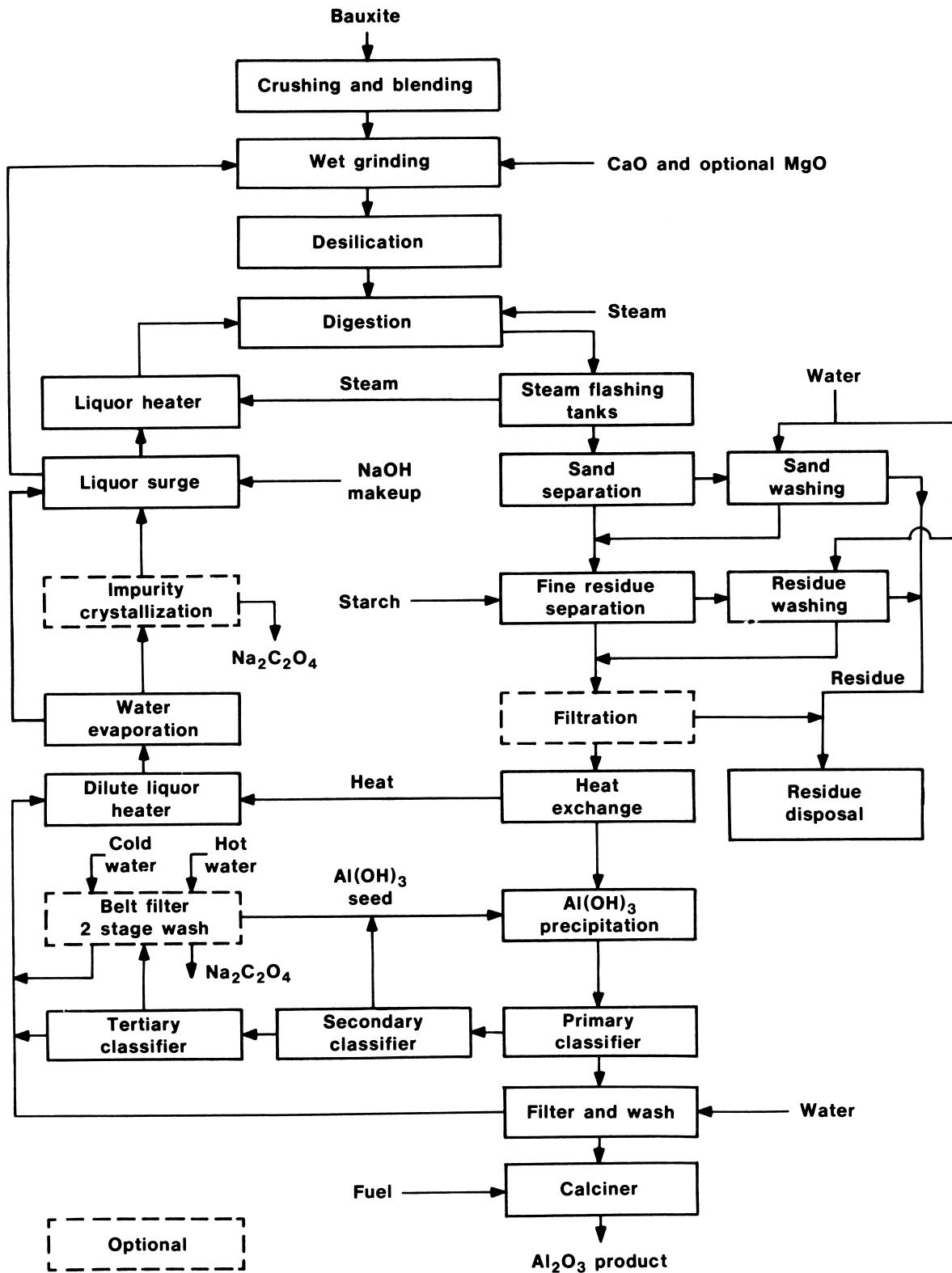
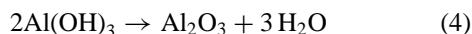


FIGURE 2 Flow diagram of Bayer process for extracting pure alumina from bauxite.

particles are filtered and washed, usually on a rotary filter. The spent liquor is reconstituted by evaporation. The washed aluminum hydroxide is calcined to alumina at 1100–1200°C.



In most modern plants the calcination is done in a fluidized-bed flash calciner. Older plants often use rotary kilns.

While most impurities are rejected with the bauxite residue, some concentrate in the liquor. Organic material enters with the bauxite, largely as humic acids and to a lesser extent as flocculating and antifoam agents. The organic material breaks down during digestion, producing intermediate compounds and ultimately sodium oxalate. Sodium carbonate also forms by reaction with CO₂ in the air. Carbonate ion (and also phosphate) is removed by precipitation with lime [Ca(OH)₂]. This restores the causticity (OH⁻ ion). The buildup of organic salts must be limited because it slows precipitation and agglomeration of aluminum hydroxide and lowers product quality. One way is to install oxalate crystallization columns following the evaporator. Sodium oxalate (and several other impurities) precipitate from the cool concentrated liquor. Another way is to wash the fine Al(OH)₃ seed with hot water to dissolve

the oxalate that coprecipitated with it. A third technique is to substitute dolime (CaO + MgO) for lime. The MgO is converted to insoluble magnesium aluminate, which has high adsorptive capacity for humic acids and is removed with the fine bauxite residue.

V. ELECTROLYTIC REDUCTION OF ALUMINA TO ALUMINUM (HALL-HÉROULT PROCESS)

Since Hall and Héroult's discovery of the process a century ago, essentially all aluminum has been produced by the electrolysis of alumina dissolved in an electrolyte of molten cryolite. A typical modern aluminum reduction cell (or pot as it is called in the industry) consists of a rectangular steel shell 9–16 m long, 3–4 m wide, and 1–1.3 m high. The shell is lined with a refractory thermal insulation that surrounds an inner lining of either carbon blocks or monolithic carbon baked in place. Carbon is the only material that can hold and withstand the combined corrosive action of aluminum and the fluoride electrolyte. Thermal insulation at the sides and ends of the cell is adjusted to provide sufficient heat loss to freeze a protective layer of electrolyte on the inner walls. Without this protective coating, aluminum depositing on the wall would

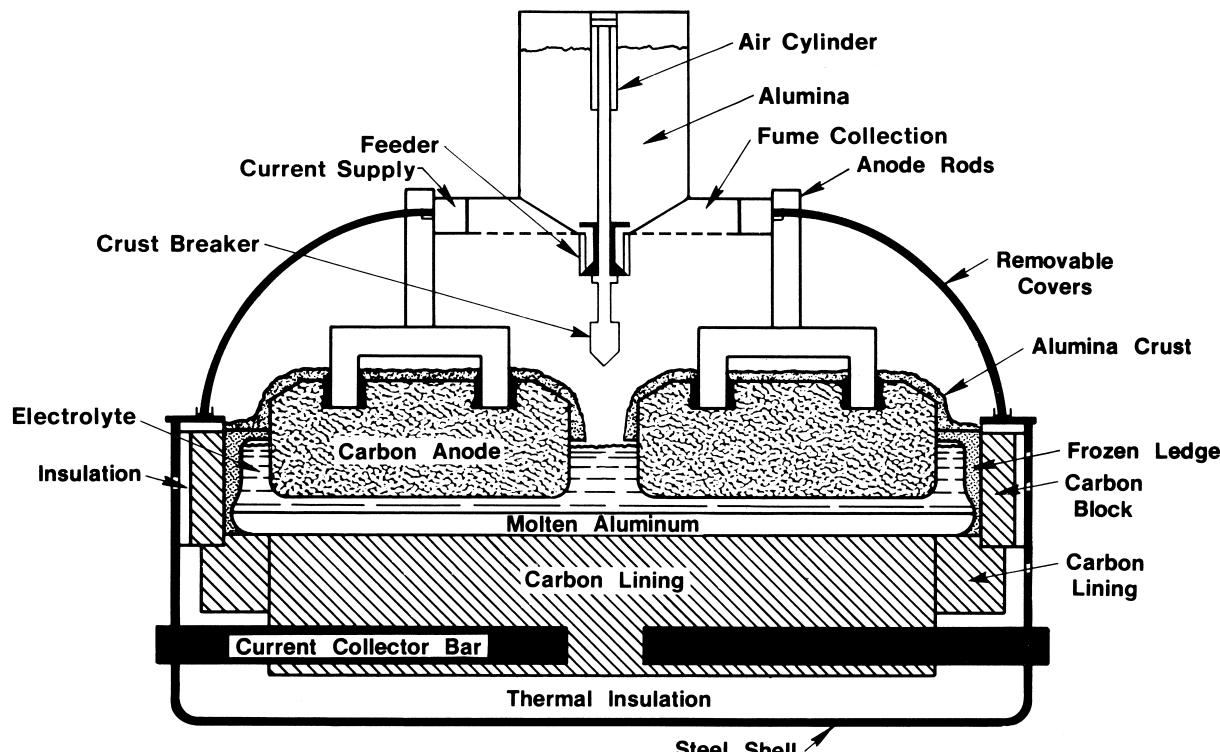


FIGURE 3 Typical Hall-Héroult cell with prebaked anodes.

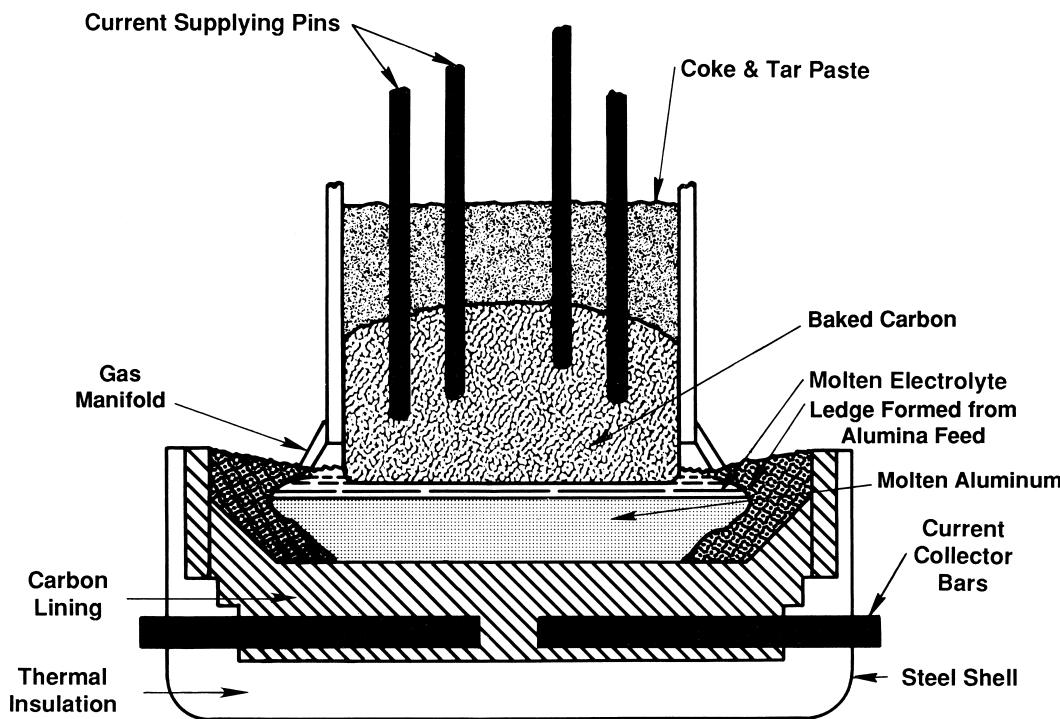


FIGURE 4 Typical Hall–Héroult cell with Söderberg anode.

produce aluminum carbide, which would then dissolve in the electrolyte, causing erosion of the wall. A crust of frozen electrolyte blanketed with alumina covers the cell. Electric current enters either through prebaked carbon anodes in cells of the type shown in Fig. 3 or through a continuously formed self-baking Söderberg anode in cells of the type shown in Fig. 4. The current then passes through 3–6 cm of electrolyte, forming CO_2 at the anode and aluminum at the cathode. Steel collector bars in the carbon lining at the bottom conduct the electric current from the cell. Today's cells range from 50 to more than 350 kA current capacity.

A. Anode Manufacture

Prebaked anodes are molded from a paste of crushed and sized petroleum coke and coal tar pitch into blocks typically 70 cm wide, 125 cm long, and 50 cm high. Petroleum coke is used because of its high purity. Metal impurities more noble than aluminum, such as iron and silicon, are deposited in the aluminum, while less noble impurities such as calcium and magnesium accumulate as fluorides in the electrolyte. The anodes are baked at temperatures between 900 and 1200°C before use. Steel stubs are inserted into sockets provided in the top surface of the anode and fixed in place by pouring cast iron or packing carbonaceous paste around them. These stubs support the anodes in the electrolyte and conduct electric current into the an-

odes. The electrical resistivity of prebaked anodes ranges from 0.005–0.006 $\Omega \text{ cm}$. Prebaked anodes are operated at current densities ranging from 0.6–1.3 A/cm^2 .

Söderberg anodes are formed continuously from a similar paste of petroleum coke and coal tar pitch. This mixture is added to the top of a rectangular steel casing typically 6–8 m long, 2 m wide, and 1 m high. The carbon paste bakes while passing through the casing, forming carbon to replace the carbon being consumed at the bottom of the electrode. The baked portion extends past the casing into the molten electrolyte. Electric current enters the anode through vertical or sloping horizontal steel pins. Periodically, the lowest pins are reset to a higher level. The electrical resistivity of Söderberg anodes is ~30% higher than prebaked anodes, while the current density employed is lower, ranging from 0.6–0.9 A/cm^2 . Although Söderberg anodes save the capital and labor required to form and bake prebake anodes, Söderberg anode cells are being phased out in the United States because they lose more voltage in the anode than prebakes and cause difficulty in collecting and disposing of the baking fumes and fluoride evolved from the electrolyte.

B. Feeding and Tapping Cells

Alumina is added to the electrolyte in appropriate increments either manually or by automatic feeders. When feeding increments are large and at the sides of the cell,

the dissolving alumina forms the protective ledge. Molten aluminum is removed from the cells, generally daily, by siphoning into a crucible. The aluminum produced is normally 99.6–99.9% pure. The typical impurities are iron, silicon, titanium, vanadium, gallium, and manganese, largely from the anode but also from impurities in the alumina.

C. Electrolyte

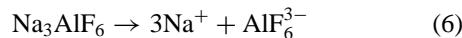
The Hall–Héroult electrolyte was chosen for its ability to dissolve alumina. It is mainly cryolite with (4–13% AlF₃) and 5–7% CaF₂) added to lower its melting point.

Cells normally are operated at temperatures ranging from 5–15°C over the liquidus temperature. Skybakmoen, Solheim and Sterten, *Metallurgical and Materials Transactions B*, **28B**, 81 (1997) give the following empirical equation for the liquidus temperature in the cryolite crystallizing range, the region of normal operation:

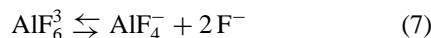
$$\begin{aligned} T_{lc} = & 1011 + 0.50(\text{AlF}_3) - 0.13(\text{AlF}_3)^{2.2} + 3.45 \\ & \times (\text{CaF}_2/[1 + 0.0173(\text{CaF}_2)]) + 0.124(\text{AlF}_3)(\text{CaF}_2) \\ & - 0.00542[(\text{AlF}_3)(\text{CaF}_2)]^{1.5} - 7.93(\text{Al}_2\text{O}_3)/ \\ & [1 + 0.0936(\text{Al}_2\text{O}_3) - 0.0017(\text{Al}_2\text{O}_3)^2] \\ & - 0.0023(\text{AlF}_3)(\text{Al}_2\text{O}_3)] - 8.9(\text{LiF})/ \\ & [1 + .0047(\text{LiF}) + 0.001(\text{AlF}_3)^2] - 3.95(\text{MgF}_2) \end{aligned} \quad (5)$$

where T_{lc} is the liquidus temperature in degrees Celsius and the components in parentheses are in weight percent.

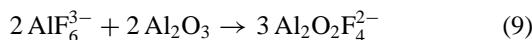
Cryolite ionizes:



The hexafluoroaluminate ion undergoes partial dissociation in the melt:



When alumina is added, oxyfluoride ions are produced:



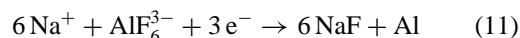
The sodium ion carries more than 95% of the current, the balance being carried primarily by the fluoride ion. The electrical conductivity of the electrolyte is given by

$$\begin{aligned} \kappa = & \exp[2.0156 - 0.0207(\% \text{Al}_2\text{O}_3) \\ & - 0.005(\% \text{CaF}_2) - 0.0166(\% \text{MgF}_2) \\ & + 0.0178(\% \text{LiF}) + 0.0063(\% \text{NaCl}) \\ & + 0.435(\% \text{NaF}/\% \text{AlF}_3) - 2068.4/T] \end{aligned} \quad (10)$$

Where κ is the conductivity in reciprocal ohm centimeters, percentages are weight percentages, and T is in Kelvins.

D. Electrode Reaction

Although sodium ions carry most of the current, aluminum is deposited at the cathode because it takes less energy to discharge aluminum than sodium. At one time it was believed that sodium was discharged and then reacted with the melt to form aluminum. Today it is generally believed that a charged transfer occurs at the cathode interface and hexafluoroaluminate ions are discharged, forming aluminum and F[−] to neutralize the charge of the current-carrying Na⁺:



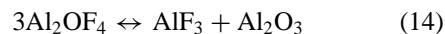
The anode process is even more complicated.

Thermodynamically, oxygen depositing on carbon at cell operating temperature should equilibrate to essentially all CO with very little CO₂. However, based on measurements of either the volume of gas produced or net carbon consumption, the primary anode product is essentially all CO₂. The reason lies in the kinetics of the reaction.

When oxygen discharges onto carbon it forms a stable C–O surface compound. The surface compound will detach slowly from the surface and produce CO gas, but at normal current densities, additional oxygen atoms react with the C–O surface compound forming CO₂ gas before it has a chance to detach. The anode reaction can be written:



This is followed by electrochemical desorption:



E. Anode Effect

When the alumina concentration falls too low, a phenomenon called an anode effect occurs. The anode dewets and voltage rises, producing sparking around the anode. This can be explained by the following mechanism. As alumina is depleted, overvoltage increases. The increase in surface tension of the bath with decreasing alumina concentration coupled with the electro-capillary effect resulting from the higher anode overvoltage reduces wetting of the anode and the bubble contact angle decreases, forming larger bubbles. At ~1.2 V anode overvoltage, sufficient fluorine activity is produced to cause fluorine to bond to the carbon. Fluorocarbon compounds have low surface

energy and promote further dewetting and growth of large bubbles. Soon a continuous film of gas lies between the bath and the anode producing the anode effect.

F. Current Efficiency

According to Faraday's law, 1 faraday (F; 26.80 A hr) should deposit 1 g equivalent (8.994 g) of aluminum. In practice only 85–96% of this amount is obtained. The loss is caused mainly by reduced species dissolving or dispersing in the electrolyte at the cathode and being transported to the vicinity of the anode, where these species are reoxidized by CO₂, forming CO and metal oxide in the electrolyte. The most likely reduced species are sodium and AlF, formed at the cathode by equilibria (13) and (14), respectively.



Equilibrium (15) predominates. Certain bath additives, particularly aluminum fluoride, and low temperatures lower the reduced-species content of the electrolyte and thereby improve current efficiency. Under certain conditions, electrical short circuiting can occur between wave crests on the aluminum pool and the anode, lowering current efficiency. Hence, considerable emphasis is placed on magnetohydrodynamic stability.

VI. ENERGY CONSIDERATIONS

Table IV gives a breakdown of the energy required to produce aluminum. In this table it was assumed that electric power for smelting was 33% hydropower at 3.6 MJ/kWh and the remainder was generated from fossil fuel at 10.9 MJ/kWh. Fabricating was assumed to use 17% hydropower, and all of the energy for mining and refining came from fossil fuels. The Bayer plants are fairly energy efficient, making extensive use of heat exchangers to recycle the heat down the temperature scale. Fluidized-bed calciners reduce energy consumption by 30% over the older rotary kiln calciners.

The melting and holding furnaces used in fabricating are only ~30% efficient, but improved firing practice, better thermal insulation, and heat recuperators are increasing this efficiency. Furnaces to bake anodes for the smelting cells use the cooling anodes to preheat combustion air and the combustion gases from the baking zone preheat incoming anodes.

While the energy used by each operation is important, it can be seen from **Table IV** that smelting consumes ~69%

TABLE IV Energy Required to Produce Aluminum (Typical Values)

Operation	Energy (MJ/kg of Al)
Mining and shipping bauxite	4.5
Mining	(1.2)
Shipping	(3.3)
Refining ore (Bayer process)	39.2
Chemicals	(2.8)
Extraction	(28.8)
Calcination	(6.0)
Shipping alumina	(1.6)
Smelting (Hall-Héroult process)	153.1
Electric power	(131.5)
Energy value of coke	(14.2)
Forming and baking anodes	(2.6)
Cathode replacement	(2.6)
Chemicals (electrolyte)	(0.3)
Environmental control	(1.9)
Fabrication ^a	26.1
Melting, holding, and casting	(6.8)
Preheating, forming, and heat treating	(19.3)
Total	222.9

^a Highly variable: depends on alloy and product.

of the total, and a major part of this is electric power. The specific power consumption in kilowatt hours of electric power used to produce a kilogram of aluminum is an important figure of merit in judging cell performance. It can be calculated,

$$\text{kWh/kgAl} = 298E_{\text{cell}}/(\% \text{CE}) \quad (17)$$

Where E_{cell} is the cell voltage and % CE the percent current efficiency. It is obvious that to minimize the specific power consumption, one must strive to maximize current efficiency and minimize cell voltage. Most cells operate at more than 90% current efficiency but at more than twice the theoretical minimum voltage. The easiest way to reduce cell voltage is to reduce anode–cathode separation, but this results in lower current efficiency and if carried to an extreme, will actually increase the specific power consumption. Calculation of voltages E_1 through E_{12} was based on the thermodynamic definition that reactions using energy are negative and reactions producing energy are positive. Engineers reverse the sign and get the individual components of voltage shown in **Figs. 5** and **6**. The thermodynamic equilibrium potential E_3 can be calculated by two different routes. One can assume that alumina is electrolytically decomposed into aluminum and oxygen.

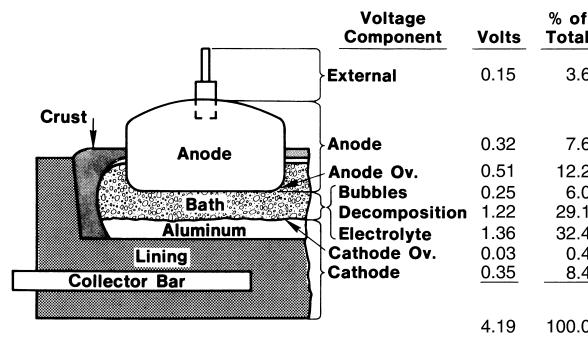
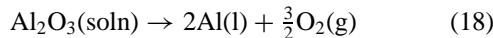


FIGURE 5 Typical distribution of voltages in a state-of-the-art aluminum production cell. Ov. indicates overvoltage.



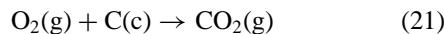
The electrical energy required for this reaction will be

$$E_1 = (-\Delta G_{(16)}^{\circ}/6F) + (RT/6F) + \ln a_{\text{Al}_2\text{O}_3} \quad (19)$$

Cell voltage measurements, made at a number of alumina concentrations, indicate that the activity of alumina can be calculated.

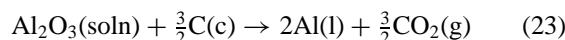
$$a_{\text{Al}_2\text{O}_3} = -0.0379R_s + 2.364R_5^2 - 2.194R_5^3 + 0.868R_5^3 \quad (20)$$

where R_s the relative saturation as the electrolyte with alumina. For an electrolyte at 960°C and 1/3 saturated, $E_1 = -2.248$. Although the cell is fed γ -alumina, the equilibrium phase at this temperature is α -alumina; therefore, α -alumina is used in calculating $\Delta G_{(16)}^{\circ}$. The oxygen deposited at the anode acts like a fuel cell, reducing the cell's voltage an amount E_2 by the reaction



$$E_2 = -\Delta G_{(19)}^{\circ}/4F \quad (22)$$

Subtracting $E_2 = -1.026$ V from E_1 makes the equilibrium potential $E_3 = -1.222$ V. Alternatively, one can assume that the cell reaction is



Then E_3 can be calculated directly,

$$E_3 = (-\Delta G_{(21)}^{\circ}/6F) - (RT/6F) \\ \times \ln [(a_{\text{Al}}^2 a_{\text{CO}_2}^{1.5}) / (a_{\text{Al}_2\text{O}_3} a_{\text{C}}^{1.5})] \quad (24)$$

These activities, except alumina, are essentially equal to unity. Equation 24 gives the same -1.222 V for E_3 . The E_4 represents anode overvoltage, normally ~ 0.5 V, calculated by

$$E_4 = (RTv/pnF) \ln(i/i^0) \quad (25)$$

where $n/\mu = 2$ electrons per oxygen molecule, p is the reaction order, ranging from 0.4 to 0.6 with carbon reactivity and porosity, R the gas constant, T the absolute temperature, F the Faraday constant, i the current density, and i^0 the reaction limiting current density, ranging from 0.0039–0.0085 A/cm² as alumina concentration varies from 2 to 8 wt %. The anode concentration overvoltage E_5 is normally a few hundredths of a volt but rises rapidly below 2 wt % alumina, leading to anode effect. One can calculate E_5 ,

$$E_5 = (RT/2F) \ln[i_c/(i_c - i)] \quad (26)$$

where i_c is the critical or concentration limiting current density, which varies from 1.0–30 with alumina

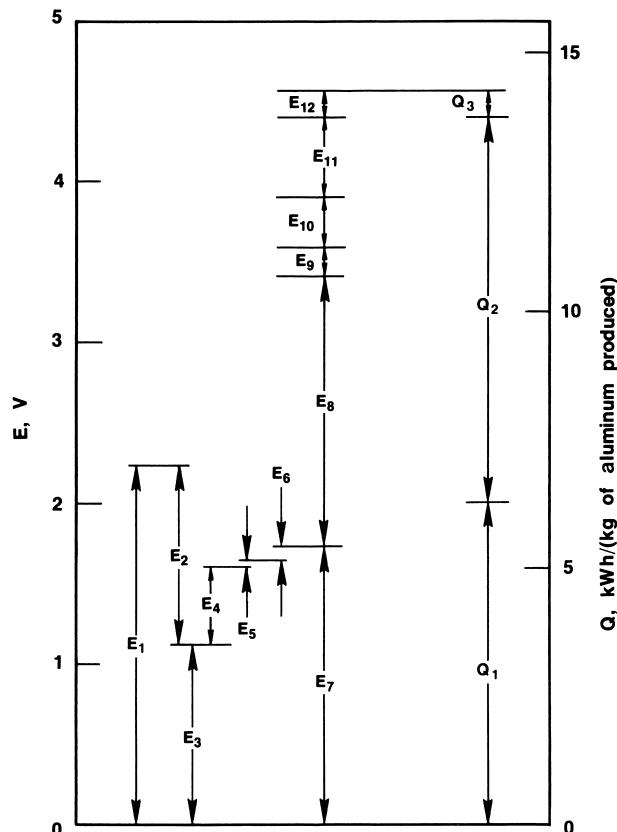


FIGURE 6 Components of cell voltage and power consumption at 95% Faradaic efficiency. E_1 , Decomposition of alumina, 2.25 V; E_2 , depolarization by carbon, -1.03 V; E_3 , equilibrium (decomposition) potential 1.22 V E_4 anode reaction overvoltage, 0.47 V; E_5 , anodic concentration overvoltage, 0.04 V; E_6 , cathodic overvoltage, 0.03 V; E_7 , counterelectromotive force of cell, 1.76 V; E_8 , ohmic losses in electrolyte, 1.33 V; E_9 , additional ohmic loss caused by bubbles, 0.25 V; E_{10} , ohmic loss in anode, 0.35 V; E_{11} , ohmic loss in cathode, 0.35 V; E_{12} , ohmic loss external to cell, 0.15 V; Q_1 , net enthalpy (energy) to produce aluminum, 6.341 kWh/kgAl; Q_2 , heat losses from cell, 6.335 kWh/kgAl; Q_3 , heat losses external to cell, 0.471 kWh/kgAl.

concentration, temperature, and size and shape of the anode. Finally, there is a cathodic overvoltage E_6 , generally about -0.03 V in industrial cells.

$$E_6 = \frac{RT[1.375 - 0.25(\% \text{NaF}/\% \text{AlF}_3)]}{1.5F} \ln \frac{i}{0.257} \quad (27)$$

These voltages sum to give E_7 , the counter-electromotive force of the cell. This is the voltage one would see if the cell current were momentarily interrupted, causing ohmic voltages to disappear. The largest voltage component is that through the electrolyte, typically -1.3 to -1.5 V excluding bubbles,

$$E_8 = IL/\kappa A \quad (28)$$

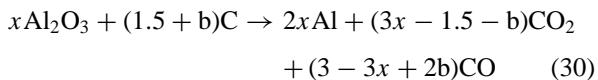
where I is the cell current in amperes, L the anode-to-cathode spacing (3–5 cm), A the cross-sectional area of the electrolyte, κ the electrical conductivity of electrolyte (1.8 – 2.3 Ω^{-1} cm $^{-1}$). Gas bubbles in the electrolyte will typically add 0.2 to 0.3 V (E_9) by reducing the cross section of electrolyte through which electric current flows,

$$E_9 = \frac{i}{\kappa} \left[\frac{d_b}{1 - \Theta} - d_b \right] \quad (29)$$

where i is current density (0.7 – 1.0 amp cm $^{-2}$), d_b is bubble layer thickness (0.48 – 0.50 cm), and Θ is the fraction of anode covered with bubbles (0.5 – 0.99).

The anode ohmic voltage E_{10} represents the potential drop through the anode, its connecting stubs, and the electrical contact between them. This loss ranges from -0.25 to -0.35 V with prebaked anodes and -0.45 to -0.55 with Söderberg anodes. The cathode voltage E_{11} includes the voltage loss in the lining of the cell, collector bars, and contact between them. It typically ranges between -0.30 and -0.55 V. There is also a voltage loss external to the cell, E_{12} , in the electrical conductors connecting one cell to the next. It typically ranges from -0.1 and -0.2 V. This voltage does not contribute heat to the cell but nevertheless represents a power loss.

The right hand side of Fig. 7 shows where the electrical energy produced by the voltages on the left is expended. The two scales correspond at 95% current efficiency. Here Q_1 represents the enthalpy of the overall cell reaction,



where x is current efficiency and b is the volume fraction of CO₂ reacting with carbon.

Note that Q_1 is greater than E_3 . A part of the joule heat is used in the reduction of alumina. The Q_2 represents heat loss from the cell and Q_3 heat generated in the bus work external to the cell. Dividing the productive energy Q_1 by the total electrical energy gives an electrical power

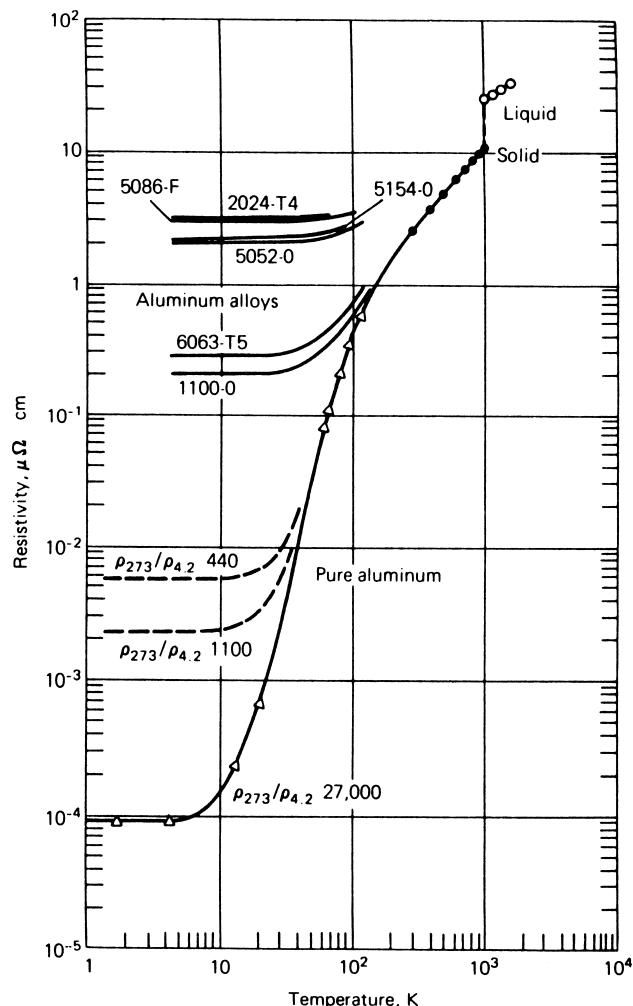


FIGURE 7 Electrical resistivity of pure aluminum and several alloys as a function of temperature. [Reprinted with permission from Hatch, J. E. (Ed.), (1984). "Aluminum, Properties and Physical Metallurgy," p. 9, American Society for Metals, Metals Park, Ohio.]

efficiency of 48%, typical for state-of-art Hall–Héroult cells.

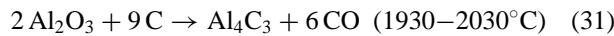
VII. ALTERNATIVE PROCESSES FOR PRODUCING ALUMINUM

While the Bayer–Hall–Héroult processes have maintained industrial dominance for over 100 years, they have several inherent disadvantages. The large number of individual cells (250–1000 in a typical plant) combined with the cost of the Bayer alumina purification plant and the carbon anode plant (or paste plant for Söderberg anodes) results in a large capital investment. The Hall–Héroult cells require expensive electric power rather than cheap thermal power. Many companies must import alumina or bauxite. The

supply of petroleum coke for anodes is limited. These disadvantages have prompted research to find alternative processes.

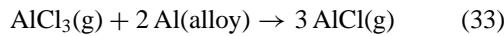
To reduce the number of units required and also save electric power, the Aluminum Company of America developed a bipolar electrode cell that electrolyzed aluminum chloride in an alkali chloride melt. The chlorine is recycled back to a chemical reactor, where chlorine, carbon, and alumina react to make more anhydrous aluminum chloride for the cell. A pilot facility using this process operated from 1976–1982. The specific electric power consumption was 30% less than their best Hall–Héroult cells. Operation of the plant was discontinued in 1982 because of marginal economics and difficulties with the chemical reactor.

Thermodynamics indicates that one should be able to reduce alumina with carbon at 2100°C. Attempts to do this, however, have produced low yields of aluminum owing to side reactions. Solid aluminum carbide (Al_4C_3) forms. Aluminum suboxide vapor (Al_2O) escapes, and the aluminum vapor produced reacts with carbon monoxide as the products leave the furnace. Yields as high as 67% can be obtained by staging the reactions as follows at two temperatures:



Adding to the furnace a second metal such as iron, silicon, tin, or copper or its oxide that is subsequently reduced to alloy with the aluminum and lower its vapor pressure provides much higher yields. To obtain pure aluminum, it is necessary to extract the aluminum from the alloy. This can be accomplished by electrochemical refining in a molten salt, but this requires extra electrical energy. Fractional crystallization can also be used, particularly well with tin–aluminum alloys because of the very low solid solubility of tin in aluminum. Professor F. J. Elliot, of the Massachusetts Institute of Technology, demonstrated in 1988 efficient carbothermic reduction of aluminum into a tin alloy in an electric furnace. Pure aluminum was then extracted from the alloy by fractional crystallization, followed by sodium treatment. Studies indicate that this process has potential for becoming competitive with the Hall–Héroult process.

Alcan in 1967 announced a process for producing aluminum by carbothermic reduction of ores to an iron–silicon–aluminum alloy and then extracting the aluminum as aluminum monochloride at 1000–1400°C:



The monohalide vapor was transported to a cooler zone (700–800°C), where the reverse of reaction (33) took

place, condensing out pure aluminum and regenerating aluminum trichloride vapor to be recycled through more impure aluminum alloy. Corrosion problems caused the project to be abandoned.

The vapor pressure of aluminum can also be lowered by operating the furnace in such a manner that the aluminum alloys with aluminum carbide as it is produced. More than 40% aluminum carbide is soluble in aluminum at 2200°C. When this alloy is tapped from the furnace and slowly cooled, the aluminum carbide crystallizes as an open lattice interspersed with pure aluminum. The pure aluminum can be separated from the aluminum carbide and the carbide residue recycled to the arc furnace. The Aluminum Company of America studied this process in the 1950s and concluded that it was uneconomic. Reynolds Metals improved upon the process in the 1980s to a point where it is close to being competitive with the Bayer–Hall–Héroult process.

Aluminum has been produced by electrolysis of aluminum sulfide dissolved in a melt of $\text{NaCl}-\text{LiCl}-\text{MgCl}_2$, producing aluminum at the cathode and sulfur at the anode. The decomposition potential is much lower than for electrolysis of alumina or aluminum chloride. Development of the process awaits an economical method for producing aluminum sulfide.

VIII. PRODUCTION OF ULTRA-HIGH-PURITY ALUMINUM

Aluminum of 99.9% purity can be produced in Hall–Héroult cells by controlling raw material impurities. This requires selection of extra-high-purity petroleum coke for the anodes and extra care in operating the cell so that steel from tools does not enter the melt and impurities are not returned to the cell through the fume treatment system.

A three-layer electrolytic refining process was developed by Hoopes of the Aluminum Company of America in the early 1900s and later improved by Pechiney and Alusuisse. Impure molten aluminum alloy, containing at least 25% copper, rests on the bottom of the cell as the anode. The electrolyte may be chlorides, fluorides, or mixed chlorides and fluorides having a density (adjusted by the addition of barium chloride or barium fluoride) less than that of the impure aluminum alloy but greater than that of pure aluminum. The pure aluminum produced floats on the electrolyte. Electrical contact to the anode is made through carbon blocks at the bottom of the cell and electrical contact to the pure aluminum at the top through graphite electrodes dipping into the aluminum. Side walls are of magnesia. Impurities more noble than aluminum are not electrolytically oxidized and accumulate in the

anode alloy. When the impurity level is sufficiently high, crystals of the impurity metals precipitate in a segregation sump at the end of the cell, where the temperature is slightly lower than in the bulk of the alloy. Impurities less noble than aluminum such as calcium, magnesium, and sodium accumulate in the electrolyte. Table V gives the characteristics of the original Hoopes cell and later improvements.

Aluminum of very high purity (99.9999%) can be produced in small quantities by zone refining. A molten zone is moved along an aluminum bar by localized heating. Impurities that lower the melting point, such as silicon and iron, concentrate at one end of the bar, while impurities that raise the melting point, such as titanium and vanadium, concentrate at the other end. After several passes, the ends are removed, leaving a middle portion of very high-purity aluminum.

Large quantities of high-purity aluminum can be produced by fractional crystallization. Impurities that raise

TABLE V Characteristics of Aluminum Refining Cells^a

Characteristic	Refining cell		
	Hoopes	Pechiney	Alusuisse
Cathode layer			
Aluminum purity (%)	99.98	99.99+	99.99+
Density (g/cm ³)	2.29	2.30	2.30
Electrolyte			
Composition (%)			
NaF	25–30	17	18
AlF ₃	30–38	23	48
CaF ₂	—	—	16
BaF ₂	30–38	—	18
BaCl ₂	—	60	—
Al ₂ O ₃	0.5–7.0	—	—
Density (g/cm ³)	2.5	2.7	2.6
Resistivity (Ω cm)	0.3	0.75–0.85	1.1
Anode layer			
Composition (%)			
Al	75	67	70
Cu	25	33	30
Density (g/cm ³)	2.8	3.14	3.05
Operating characteristic			
For a cell of amperage	20,000	25,000	14,000
Volts	5–7	6.9	5–6
Current density (A/cm ²)	0.95	0.40	0.36
Current efficiency (%)	90–98	96–98	92–95
Operating temperature (°C)	950–1000	750	740

^a Data from T. G. Pearson, "The Chemical Background of the Aluminum Industry," Royal Institute of Chemistry, Monograph No. 3, London 1955 and L. Evans and W. B. C. Perrycoste, British Intelligence Objectives Subcommittee Final Report No. 1757, Item 21 (1946).

the melting point and react with boron to form insoluble borides are chemically precipitated in a holding furnace. After the melt is transferred to another furnace, crystals of high-purity aluminum are formed by cooling the surface of the molten aluminum. When the furnace is nearly filled with crystals, the remaining liquid is drained. Then the crystals are remelted, yielding 99.995% or greater pure aluminum.

IX. PROPERTIES OF ALUMINUM

A. Physical Properties

Aluminum is a silvery white metallic element in Group 3A of the periodic chart (Group 13 by the new IUPAC notation), atomic number 13, atomic weight 26.9815. Some of the physical properties of 99.996% pure aluminum are given in Table VI. Many of the properties vary significantly with purity or alloying. In fact, the main purpose of alloying is to increase the hardness and strength of aluminum. The electrical resistivity of pure aluminum and several alloys is given in Fig. 7. Aluminum becomes superconducting at temperatures below 1.2 K.

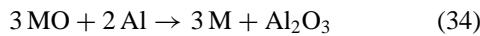
B. Inorganic Compounds

The aluminum atom has an electronic configuration of $1s^2 2s^2 2p^6 3s^2 3p^1$ and has a valence of +3 in all compounds with the exception of a few high-temperature monovalent and divalent gaseous species. High-purity aluminum is resistant to attack by most acids. It is used for storage of nitric acid, concentrated sulfuric acid, organic acids, and other chemical reagents. Because of its amphoteric nature, aluminum is attacked rapidly by alkali hydroxide solutions with the evolution of hydrogen and the formation of soluble aluminates. Aluminum reacts vigorously with oxygen-free fluorine, chlorine, bromine, and iodine, forming trihalides. It reacts with chlorinated hydrocarbons in the presence of water. Above 800°C aluminum reacts with trihalides to form gaseous aluminum monohalides. On cooling, the monohalides disproportionate to the normal trivalent compounds and aluminum. Neither aluminum hydroxide nor aluminum chloride solutions ionize appreciably but behave in some respect as covalent compounds. The aluminum ion has a coordination number of 6, and in aqueous solution it binds six water molecules as $Al(H_2O)_6^{3+}$.

Aluminum reacts at elevated temperatures with nitrogen, carbon, sulfur, and phosphorus to form aluminum nitride, aluminum carbide, aluminum sulfide, and aluminum phosphide, respectively. Aluminum powder dispersed in air can explode if ignited. Aluminum reduces many oxides:

TABLE VI Physical Properties of Aluminum

Property	Value	Property	Value as function of temperature ^b
Atomic number	13	Heat capacity (300–933.5 K)	$418.61 + 2.9338T - 6.0839E - 3T^2 + 6.098E - 6T^3 - 2.10E - 9T^4 \text{ kJ/(kg K)}$
Atomic weight	26.9815		
Crystal structure (4–933 K)	Face-centered cubic		
Coordination number	12		
Atoms per unit cell	4		
Lattice constant (298 K)	0.40496 nm	Heat capacity (0–200 K)	$-5.342E - 4 + 5.1339T + 1.0632E - 3T^2 - 4.981E - 5T^3 + 7.50E - 8T^4 \text{ kJ/(kg K)}$
Melting point	933.52 K		
Boiling point	2767 K		
Heat capacity (298 K)	0.8972 kJ/kg/K		
Heat capacity (liquid)	1.2735 kJ/kg/K		
Heat of fusion	397 J/g	Density (solid) (298–933.5 K)	$2852 - 0.888T + 0.00152T^2 - 9.85E - 7T^3 \text{ kg/m}^3$
Heat of vaporization	10,777 J/g		
Tensile strength	50 MPa		
Young's modulus	65,000 MPa	Density (liquid) (933.5–1173 K)	$2546 - 0.1659T - 2.91E - 8T^3 \text{ kg/m}^3$
Density (298 K)	2.698 g/cm ³		
Electrical resistivity (293 K)	2.655 $\mu\Omega$ cm	Thermal conductivity (200–933.5 K)	$2.222 + 7.018E - 4T^2 - 2.539E - 9T^3 + 1.662E - 12T^4 - \text{W/(cm K)}$
Thermal conductivity (298 K)	2.37 W/(cm K)		
Electrode potential versus SHE, ^a low pH (25°C)	-1.660 V	Thermal conductivity (933.5–2200 K)	$0.681 + 4.589E - 7T^2 - 2.466E - 10T^3 + 3.722E - 14T^4$
Electrode potential versus SHE, ^a high pH (25°C)	-1.706 V		
Magnetic susceptibility (298 K)	0.016 m ⁻³ g atom ⁻¹	Electrical resistivity (273–933.5 K)	$0.3992 + 0.01049T + 2.039E - 9T^3 \mu\Omega \text{ cm}$
Reflectivity at 0.9 to 12 μm	>0.9		
Reflectivity at 0.2 μm	0.7		
Thermal neutron cross section	0.232 \pm 0.003 barn	Electrical resistivity (933.5–1473 K)	$10.398 + 0.01475T \mu\Omega \text{ cm}$
		Surface tension (933.5–1273 K)	1010 – 0.152T dynes/cm
		Viscosity (933.5–1273 K)	0.1492 exp(1985/T) mPas
		Linear expansion (213–373 K)	$L_{273} \text{ K} [1 + (22.17t + 0.012t^2)10^{-6}]$
		Vapor pressure (1000–2600 K)	$\exp[17.336 - (34599/T) - (1.4421E - 6/T^2)] \text{ kPa}$

^a SHE, Standard hydrogen electrode.^b T, Temperature (K); t, T – 273.2°C; L₂₇₃ K, length at 273.2 K, or 0°C.

Aluminothermic reduction is used to manufacture certain metals and alloys and in thermite welding. These reactions, although very exothermic, require a high temperature such as that produced by burning magnesium ribbon to ignite them.

Hydrogen is the only gas known to have appreciable solubility in solid or molten aluminum. Hydrogen can be introduced into liquid aluminum during melting by reaction with the moisture in the furnace atmosphere, by moisture trapped in the oxide film on the solid aluminum, or by

moisture in the furnace refractories. A 20-fold reduction in solubility occurs on freezing, producing porosity in the solid if hydrogen in the molten aluminum is not reduced to a low level, as by fluxing.

C. Organic Compounds

Aluminum does not react with most hydrocarbons, saturated or unsaturated, aliphatic or aromatic. Most pure halogenated hydrocarbons and other halogenated organic compounds do not react with aluminum, except at elevated temperatures or in the presence of water, which

causes hydrolysis. However, some halogenated organic compounds may react violently with aluminum. Before aluminum is used with any halogenated compound, their compatibility should be established. Aluminum shows very good resistance to alcohols that contain a trace of water. Anhydrous alcohols react with aluminum, forming alcoholates. Aluminum does not react with aldehydes, ketones, or quinones. There are many aluminum metalloc-organic compounds of the type AlXR, where X may be oxygen, nitrogen, or sulfur and R an organic radical. Alcoholates or alkoxides are compounds of this type, with R being an alcohol. Certain polymers include compounds with aluminum–carbon bonds. They may be linear or cross-linked vinyl-, divinyl-, and trivinylaluminum halides. These compounds polymerize rapidly to compounds whose structures are not completely understood. Aluminum alkyls also contain aluminum–carbon bonds. However, they are not polymers but are bridge compounds of differing complexity. Alkylaluminum compounds are used as catalysts in the preparation of oriented crystalline polyolefins. These, Ziegler–Natta catalysts, have made it possible to produce elastomers that are essentially identical in structure to natural rubber.

D. Effect of Alloying

There are two general types of aluminum alloys: non-heat-treatable, work-hardening alloys and heat-treatable alloys. The non-heat-treatable alloys are typified by the solid-

solution-hardened magnesium alloys and the dispersion-hardened manganese alloys. Manganese does not remain in solid solution on cooling but precipitates as a finely dispersed aluminum–manganese constituent. Strain hardening, resulting from cold work, is caused principally by lattice distortions and associated dislocation interactions. The dislocation density can be reduced by heating the metal. This annealing heat treatment also causes recrystallization and growth of more nearly equiaxed grains. The new structure is largely free of strain, has substantially fewer dislocations within the grains, and has essentially no concentration gradients at grain boundaries. A certain amount of room-temperature annealing called age softening occurs. Because age softening complicates the specification of minimum properties for strain-hardened alloys, it is common practice to accelerate age softening and increase ductility by heating the alloy briefly to 120–175°C. Aluminum of very high purity can recrystallize at room temperature. Dispersed phases in aluminum may act either to accelerate or to retard recrystallization depending on their size, interparticle spacing, and stability at the annealing temperature.

Heat-treatable alloys contain alloying elements that form on solidification constituents (such as CuAl₂) in excess of the equilibrium solid solubility at room temperature or moderately above room temperature. These constituents are put into solid solution by raising the temperature above the solvus line of the phase diagram (see Fig. 8 for an example), holding this temperature long enough to

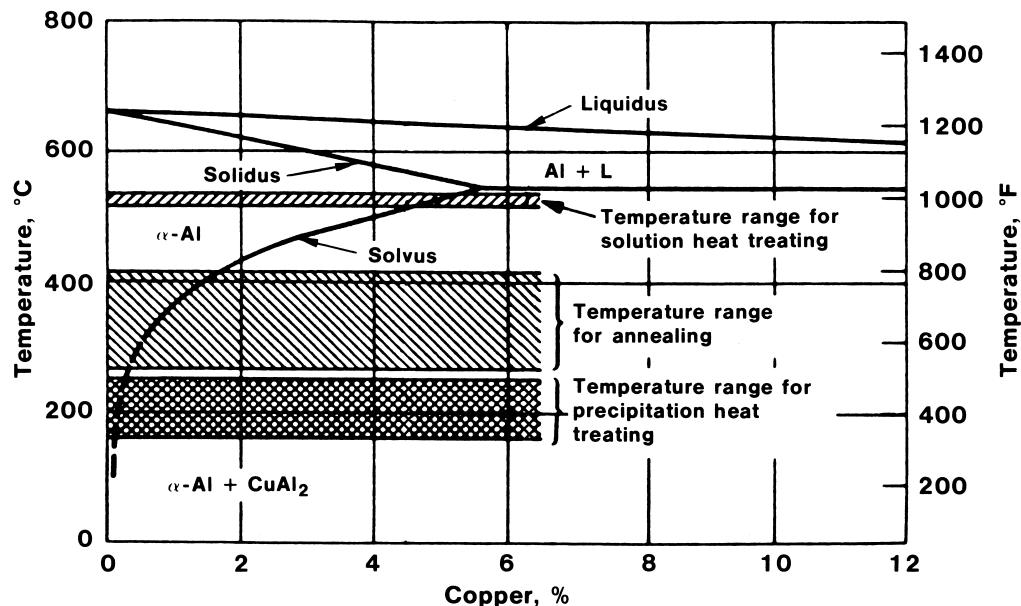


FIGURE 8 Phase diagram for an aluminum-copper system showing heat-treating temperature ranges. [Reprinted with permission from Hatch, J. E. (ed.), (1984). "Aluminum, Properties and Physical Metallurgy," p. 135, American Society for Metals, Metals Park, Ohio.]

allow the alloying elements to enter into solid solution, then quenching rapidly enough to hold these elements in solid solution. Most alloys exhibit natural aging, in which the elements in supersaturated solid solution precipitate. This may require only a few days to several years, depending on the alloy. Precipitation can be accelerated by heating above room temperature in an operation referred to as artificial aging or precipitation heat treatment. The initial stage is a redistribution of solute atoms within the solid-solution lattice to form clusters called Guinier-Preston, or GP, zones that are considerably enriched in solute. This local segregation of solute atoms produces a distortion of the lattice both within the zones and extending several atoms into the matrix. These zones interfere with the motion of dislocations as they cut across the GP zones, producing a hardening effect. As artificial aging increases, GP zones are transformed into particles different in structure from both the solid solution and the equilibrium phase. The transition phase has a specific crystallographic orientation relative to the solid solution such that the two phases remain coherent on certain planes by adaptation of the solid-solution matrix. This produces an elastic strain, giving a hardening effect that increases as the size of the particles increases. As the transition precipitate continues to grow, the coherency strain eventually exceeds the interfacial bond and the precipitant is transformed into its equilibrium form. This loss of lattice strain results in some loss in strength (overaging), but the precipitant still produces some hardening by causing dislocations to loop around the particle.

E. Corrosion Resistance

Most aluminum alloys have good corrosion resistance resulting from a thin, compact, adherent oxide film that forms on the surface. Without this protective film, aluminum is a reactive metal with poor corrosion resistance. Whenever a freshly created aluminum surface is exposed to air or water at room temperature, an oxide film forms immediately and grows to a thickness of ~ 5 nm (50 \AA) in air and to a somewhat greater thickness in water. The film thickness increases at elevated temperature. Aluminum oxide dissolves in strong acids and alkalis. When the film is removed, aluminum corrodes rapidly by uniform dissolution. Most constituents formed by the interaction of aluminum and the alloying element have electrode potentials that differ from pure aluminum and from the solid solution in which the constituent exists (Table VII). Since the oxide film over these constituents may be weak, the particle may promote electrochemical attack.

Breakdown of the oxide coating often produces pitting corrosion. Iron and copper as impurities or as alloying agents are prone to promote pitting. Pure alu-

TABLE VII Corrosion Potentials^a

Constituent	Potential (mV)
Mg	-880
Mg ₂ Al ₃	-390
SS Al + 4% MgZn ₂	-200
SS Al + 4% Zn	-200
MgZn ₂	-220
Zn	-170
CuMgAl ₂	-150
SS Al + 1% Zn	-110
SS Al + 7% Mg	-40
SS Al + 5% Mg	-20
MnAl ₆	0
Al (99.95%)	0
SS Al + 1% Mg ₂ Si	+20
SS Al + 1% Si	+40
SS Al + 2% Cu	+100
CuAl ₂	+120
SS Al + 4% Cu	+160
Steel	+270
FeAl ₃	+290
Pb	+300
NiAl ₃	+330
Si	+590
Cu	+650
Ag	+770
Ni	+780

^a Corrosion potential relative to 99.95% pure aluminum measured in an aqueous solution of 53 g per liter NaCl + 3 g per liter H₂O₂ at 298 K. SS indicates solid solution.

minum, aluminum-manganese, aluminum-magnesium, and aluminum-magnesium-silicon alloys have excellent resistance to corrosion in most weathering exposures, including industrial and seacoast atmospheres. The high-strength aluminum-copper (2000 series) and aluminum-copper-magnesium-zinc (7000 series) alloys have significantly lower corrosion resistance. These alloys are often protected against corrosion by cladding them either with pure aluminum, a low magnesium-silicon alloy, or an alloy containing 1% zinc. The cladding may be on one or both sides, depending on use, and represents 2.5–5% of the total thickness. In addition to its excellent inherent corrosion resistance, the cladding provides galvanic protection of the core material.

Intergranular corrosion is caused by selective electrochemical attack at grain boundaries when the microconstituents there have a different electrochemical potential than the adjacent depleted solid solution. Susceptibility to intergranular corrosion is controlled by fabrication practices, particularly heat treatments, that affect the amount,

TABLE VIII Nominal Compositions of Typical Aluminum Foundry Alloys^a

Aluminum Association number	Casting process ^b	Alloying elements (%)				Applications
		Si	Cu	Mg	Others ^c	
208.0	S	3.0	4.0			General purpose
213.0	P	2.0	7.0			Cylinder heads, timing gears
242.0	S, P		4.0	1.5	2.0 Ni	Cylinder heads, pistons
295.0	S	1.1	4.5			General purpose
B295.0	P	2.5	4.5			General purpose
308.0	P	5.5	4.5			General purpose
319.0	S, P	6.0	3.5			Engine parts, piano plates
A332.0	P	12.0	1.0	1.0	2.5 Ni	Pistons, sheaves
F332.0	P	9.5	3.0	1.0		Pistons, elevated temperatures
333.0	P	9.0	3.5	0.3		Engine parts, meter housings
355.0	S, P	5.0	1.3	0.5		General—high strength, pressure tightness
356.0	S, P	7.0		0.3		Intricate castings—good strength, ductility
360.0	D	9.5		0.5	2.0 Fe max	Marine parts, general purpose
380.0	D	8.5	3.5		2.5 Fe max	General purpose
A413.0	D	12.0				Large intricate parts
443.0	D	5.3			2.0 Fe max	Carburetors, fittings, cooking utensils
B443.0	S,P	5.3			0.8 Fe max	General purpose
514.0	S			4.0		Hardware, tire molds, cooking utensils
520.0	S			10.0		Aircraft fittings
A712.0	S	0.5	0.7	6.5 Zn		General purpose

^a From "Aluminum Standards and Data, 1982." Aluminum Association, Washington, D.C., 1982.

^b Sand cast; P, permanent mold cast; D, pressure die cast.

^c Aluminum and impurities constitute remainder.

size, and distribution of the precipitate. It is desirable to have a fine, distributed precipitation throughout the grain structure. Precipitants that have the same potential as the matrix, such as MnAl₆, do not cause intergranular corrosion.

Exfoliation results from layered or lamellar corrosion. It usually occurs along grain boundaries and hence is a specialized case of intergranular corrosion. Exfoliation can be caused by striations of insoluble constituents resulting from severe cold working.

The combined action of a tensile stress and a corrosive environment can produce stress corrosion cracking. Intergranular corrosion produces a notch. Stress causes the notch to open, allowing corrosion to progress until the part is so weakened that it eventually fails mechanically.

X. ALLOYS OF ALUMINUM

A. Foundry Alloys

Table VIII gives the composition of common foundry alloys. Unalloyed aluminum does not have the necessary mechanical properties or casting characteristics for foundry use. Silicon is the most widely used alloy-

ing element. It improves fluidity, increases resistance to hot cracking, and improves pressure tightness. Binary aluminum–silicon alloys, however, have relatively low strengths. Hence, other elements such as copper and magnesium are added to obtain higher strengths through heat treatment. A small amount of sodium or strontium is frequently added to modify the aluminum–silicon eutectic to give increased ductility. Small additions of titanium or titanium and boron act as grain refiners to increase resistance to hot cracking and improve tensile properties. Alloys for die casting have a high iron content to prevent welding to the dies. The F temper designates the properties of alloys as casted, and the T temper designates the properties after heat treatment.

B. Wrought Alloys

The composition of commercial wrought alloys is given in Table IX. The 1000 series comprises commercially pure aluminum (99% pure or higher), although 0.12% copper is added to 1100 alloy. Iron and silicon are the major impurities. These compositions are characterized by excellent corrosion resistance, high thermal and electrical conductivity, low mechanical properties, and excellent

TABLE IX Nominal Compositions of Typical Wrought Aluminum Alloys^a

Aluminum Association number	Alloy elements (%)						Applications
	Si	Cu	Mn	Mg	Cr	Zn	
1100		0.12				1.0 (Si + Fe) max	Sheet metal work, pots, pans
1350						99.5 Al min	Electrical conductor
2014	0.8	4.4	0.8	0.5			Aircraft structures, truck frames
2024		4.4	0.6	1.5			Aircraft structures, truck wheels
2036		2.6	0.25	0.45			Automotive sheet
2090		2.7				2.2 Li, 0.12 Zr	Aircraft structures
2091		2.2		1.5		2.0 Li, 0.08 Zr	Aircraft structures
2219		6.3	0.3			0.1 V, 0.18 Zr, 0.06 Ti	Aerospace structures, welded
3003		0.12	1.2				General purpose, cooking utensils
3004			1.2	1.0			General purpose, can sheet
3105			0.6	0.5			Building products
5052				2.5	0.25		General purpose
5083			0.7	4.4	0.15		Unfired pressure vessels
5182			0.35	4.5			Can sheet, automotive sheet
5252			2.5				Automotive bright trim
6009	0.8	0.33	0.33	0.5			Automotive sheet
6010	1.0	0.33	0.33	0.8			Automotive sheet
6061	0.6	0.28		1.0	0.20		General purpose, structures
6063	0.4			0.7			General purpose, extrusions
6201	0.7			0.8			Electrical conductor
7005			0.45	1.4	0.13	4.5	0.14 Zr, 0.04 Ti
7075		1.6		2.5	0.23	5.6	Aircraft structures
8090		1.3		0.9		2.4 Li, 0.12 Zr	Aerospace structures

^a From "Aluminum Standards and Data, 1982." The Aluminum Association, Washington, D.C., 1982.

^b Aluminum and impurities constitute remainder.

workability. A moderate increase in strength can be obtained by strain hardening. These alloys have many applications, especially in the electrical and chemical fields.

Copper is the principal alloying element in the 2000 series. These alloys require solution heat treatment to obtain optimum properties. In the heat-treated condition, mechanical properties are similar to and may even exceed those of mild steel. Sometimes artificial aging is employed to improve the mechanical properties even more. While this treatment has little effect on tensile strength, the yield strength is materially increased with an attendant loss in elongation. These alloys have less corrosion resistance than most other aluminum alloys and under certain conditions may be subject to intergranular corrosion. Therefore, the 2000 series alloys in sheet form are usually clad with either high-purity aluminum or a magnesium–silicon alloy of the 6000 series to provide galvanic protection of the core material. Alloy 2024 is probably the best known, having been the main World War II aircraft alloy.

Manganese is the major alloying element of the 3000 series. These alloys generally are not heat treatable but

are appreciably stronger than the 1000 series. Alloy 3003 is a popular, general purpose alloy for moderate-strength applications where good workability and weldability are required.

The major alloying element in the 4000 series is silicon. Silicon can produce a substantial lowering of the melting point without causing brittleness in the resulting alloys. These alloys are used, therefore, as welding wire and as brazing alloys, where a melting point lower than that of the parent metal is required. Although most of the alloys in this series are not heat treatable, when they are used for welding or brazing heat-treatable alloys, they dissolve enough constituents from the base alloy to respond later in a limited manner to heat treatment. Alloys containing an appreciable amount of silicon develop an attractive dark gray color when anodized. Hence, these alloys are frequently used for architectural applications.

Magnesium is the major alloying element in the 5000 series. It is one of the most effective and most widely used alloying elements for aluminum. Combination with manganese results in a moderate-to-high-strength

non-heat-treatable alloy. Not only is magnesium 50–60% more effective than manganese as a solid-solution hardener; it can be added in considerably greater quantities. These alloys possess good welding characteristics and good corrosion resistance even in marine atmosphere. Limitations, however, must be placed on the amount of cold work and operating temperatures for the alloys of higher magnesium content to avoid susceptibility to stress corrosion cracking.

Magnesium and silicon are the major alloying elements in the 6000 series alloys. These alloys are heat treatable, Mg_2Si being the strengthening precipitate. The popular 6061 alloy is one of the most versatile of the heat-treatable alloys. Although not as strong as most of the 2000 or 7000 series alloys, the magnesium silicide alloys combine good formability with good corrosion resistance.

Zinc is the major alloying element in the 7000 series group. When zinc is coupled with a small amount of magnesium, the result is a heat-treatable alloy with very high strength. Other elements such as copper and chromium are generally added in small quantities. The outstanding alloy of this group is 7075. It is one of the highest strength alloys available and is used in airframe structures for highly stressed parts.

Aluminum–lithium alloys such as 2090, 2091, and 8090 will probably become the aerospace alloys of the future. Lithium both lowers the density and increases the elastic modulus. Moreover, these alloys are heat treatable. Addition of magnesium to the binary aluminum–lithium alloys reduces the density and increases the strength without affecting the improved modulus. Copper also increases the strength of the alloy, but if added in significant quantities it increases the density. Strengthening of the binary alloy comes from an Al_3Li precipitate. In the magnesium ternary, $LiMgAl_2$, and Al_3Li , precipitates are formed in addition to solid-solution hardening.

C. Temper Designations

A system of letters and numbers has been chosen by the Aluminum Association to designate the various aluminum alloy tempers. Temper designations follow the alloy designation separated by a hyphen. The F temper indicates as fabricated, with no special control over thermal conditions or strain hardening; O indicates annealed; H indicates strain hardened and is applied to wrought products only; W indicates solution heat treated, an unstable temper; T indicates thermally treated to produce a stable temper. The H tempers are subdivided into H1, strain hardened only; H2, strain hardened and partially annealed; H3, strain hardened and stabilized. The second digit of the H temper indicates the degree of cold work from 0 for none to 8 for 75% reduction. The T tempers are subdivided into T1, cooled from an elevated temperature shaping process

and naturally aged; T2, cooled from an elevated temperature shaping process, cold worked, and naturally aged; T3, solution heat treated, cold worked, and naturally aged; T4, solution heat treated and naturally aged; T5, cooled from an elevated temperature shaping process and artificially aged; T6, solution heat treated and artificially aged; T7, solution heat treated and overaged; T8, solution heat treated, cold worked, and artificially aged; T9, solution heat treated, artificially aged, and then cold worked; T10, cooled from an elevated temperature shaping process, cold worked, and then artificially aged. Typical properties of some foundry alloys and wrought alloys are given in Tables X and XI, respectively.

TABLE X Typical Mechanical Properties of Aluminum Foundry Alloys^a

Alloy and temper	Tensile strength ^b (MPa)	Yield strength ^b (MPa)	Elongation (%)	Brinell hardness ^b
Sand castings				
208.0-F	145	95	2.5	55
242.0-T21	185	125	1	70
242.0-T77	205	160	2	75
295.0-T4	220	110	9	60
295.0-T6	250	165	5	75
319.0-F	185	125	2	70
319.0-T6	250	165	2	80
355.0-T6	240	170	3	80
356.0-T6	225	165	3	70
B443.0-F	130	55	8	40
514.0-F	170	85	9	50
520.0-T4	330	180	16	75
A712.0-T5	240	170	5	75
Permanent mold castings				
242.0-T61	325	290	0.5	110
308.0-F	195	110	2	70
319.0-F	235	130	2.5	85
319.0-T6	275	185	3	95
F332.0-T5	250	195	1	105
355.0-T6	290	185	4	90
356.0-T6	260	185	5	80
B443.0-F	160	60	10	45
Die castings				
360.0	305	170	3	75
380.0	315	160	3	80
A413.0	290	130	3	80
443.0	225	95	9	50

^a From "Aluminum Standards and Data, 1982," the Aluminum Association, Washington, D.C., 1982.

^b These data are intended as a basis for comparing alloys and tempers and should not be used for design purposes. Properties may vary with method of manufacture, product size, and form.

TABLE XI Typical Mechanical Properties of Wrought Aluminum Alloys^a

Alloy and temper	Tensile strength ^b (MPa)	Yield strength ^b (MPa)	Elongation ^b (%)	Brinell hardness ^b	Elastic modulus ^b (MPa × 10 ³)
1100-O	90	35	35	23	69
1100-H14	125	115	9	32	69
1100-H18	165	150	5	44	69
1350-O	85	30	—	—	69
1350-H19	185	165	—	—	69
2014-T6	485	415	13	135	73
2024-T3	485	345	18	120	73
2024-T81	485	450	7	125	73
2090-T83	560	520	7	—	80
2091-T84	420	315	18	—	79
2219-T62	415	290	10	110	73
2219-T81	455	350	10	123	73
3003-O	110	40	30	28	69
3003-H14	150	145	8	40	69
3003-H18	200	185	4	55	69
3004-O	180	70	20	45	69
3004-H34	240	200	9	63	69
3004-H38	285	250	5	77	69
3105-O	115	55	24	—	69
3105-H25	180	160	8	—	69
5052-O	195	90	25	47	70
5052-H34	260	215	10	68	70
5052-H38	290	255	7	77	70
5083-O	290	145	22	65	71
5083-H321	315	225	16	82	71
5182-O	270	130	23	64	71
6009-T4	235	130	24	63	69
6010-T4	295	185	23	67	69
6061-O	125	55	25	30	69
6061-T4	240	145	22	65	69
6061-T6	310	275	12	95	69
6063-T6	240	215	12	73	69
7005-T53	395	345	15	105	70
7075-T6	570	505	11	150	72
7075-T76	525	450	11	—	72
7075-T73	505	415	11	—	72
8090-T81	460	360	7	—	81

^a From "Aluminum Standards and Data, 1989." The Aluminum Association, Washington, D.C., 1989.

^b These data are intended as a basis for comparing alloys and tempers and should not be used for design purposes.

Properties may vary with manufacture process, product size, and form.

XI. FORMING AND SHAPING OPERATIONS

A. Casting

Aluminum tapped from the smelting cells is generally transported in the molten state to a holding furnace supplying an on-site ingot-casting facility. Sometimes alu-

minum is transported distances of up to 200 km on trucks or railroad cars equipped with well-insulated containers for keeping the aluminum molten until it is alloyed and cast at a remote foundry. This conserves the energy required for remelting. Aluminum can be cast by all known methods including the use of sand molds and permanent molds, as well as die casting. When foundries are not within molten metal shipping distance, the aluminum is

cast into either unalloyed or alloyed pig or remelt ingots in stackable 22-, 340-, or 680-kg sizes, depending on the customer's preference.

Aluminum to be cast into ingots is blended with remelted scrap in the holding furnace, alloyed, and analyzed before being cast. With automated equipment, a complete metal analysis can be done in 3–5 min. The metal then passes through an in-line fluxing and filtering unit(s) as it flows to the caster. Fluxing with a gas mixture containing chlorine or fluorine effectively removes harmful sodium, lithium, and calcium and lowers the hydrogen content to an acceptable level. Filtering removes particulate matter. Ingots for rolling can weigh more than 17 metric tons. Most ingots for rolling and extrusion and some remelt ingots are produced by continuous or semi-continuous direct chill (DC) casting whereby solidified aluminum is continuously removed from the bottom of an open-ended mold as molten aluminum flows into the top. Water sprays remove heat from the exiting aluminum. In the latest technology, an alternating current magnetic field holds the freezing shells of aluminum away from the sides of the mold to produce a better surface on the ingot. An alternate process uses a stream of compressed air to prevent the freezing metal from contacting the mold. The objective is to avoid scalping, which is normally required on DC-cast wrought alloy ingots before they are extruded or rolled. The surface of the ingot is machined smooth in the scalping operation.

B. Fabrication

In producing wrought material, the cast ingot (or billet) is heated to 720 K or higher, depending on the alloy, to improve the chemical homogeneity. The initial reduction in thickness in producing aluminum sheet is achieved by hot rolling. Subsequent reductions may be hot or cold depending on the properties required in the final product. Annealing may be required during fabrication to restore the ductility lost during cold forming. Most sheet is produced as coiled strip for ease and economy of handling.

In extrusion a heated billet is forced to flow through a die opening much as toothpaste exists from a toothpaste tube. Many shapes, including hollows, can be produced that would otherwise be difficult to obtain. Extruded rod is frequently the stock for wire drawing. forgings are produced by hammering ingots into simple shapes on flat dies or into complex shapes in cavity dies.

Aluminum and aluminum alloy powders are produced by atomizing (spraying the molten metal into a cool gas, usually air). The powders are used as pigment in paints, in powder metallurgy, and as a component in explosives and rocket fuels.

XII. FINISHES FOR ALUMINUM

Aluminum can be finished in more ways than any other common metal. Processes for producing finishes, both decorative and protective, include anodizing, chemical conversion coating, electroplating, porcelain enameling, painting, etching, and mechanical texturing. Anodizing is an electrochemical process in which the natural oxide film is increased in thickness by making the aluminum the anode in an electrolytic cell containing a suitable electrolyte. Sulfuric acid, the most common electrolyte, produces a clear oxide coating on pure aluminum, high-purity aluminum–magnesium, and aluminum–magnesium silicide alloys. As the purity of the metal decreases, constituents and their reaction products collect in the oxide film, decreasing its transparency and giving the film a gray color. Automotive trim is made from high-purity aluminum to obtain maximum brightness and specularity. Colored oxide films are produced by anodizing in a mixed-acid electrolyte containing certain organic sulfonic acids and sulfuric acid. The colors produced range from pale yellow to dark broze and black. These coatings are widely employed in architectural applications and usually have a coating between 0.02 and 0.03 mm thick. Mechanical finishing techniques such as wire scratch brushing, sand blasting, and hammering are used often in combination with anodizing for special effects.

Chemical conversion treatments react with the aluminum to produce a chemical film that is integral with the metal surface. Carbonate, chromate, and phosphate films have excellent adhesion qualities and can be used as an undercoat before painting. Zinc or tin emulsion coatings are used as a base for subsequent electroplating, which can apply any of the commonly electroplated metals.

Porcelain enameling requires the use of frits with melting temperatures below 550°C. The enamels are applied over a chemical conversion coating compatible with the frit. Wrought alloys 1100, 3003, and 6061 and casting alloy 356 give best spall resistance.

XIII. USES OF ALUMINUM

The sizes of the various U.S. markets for aluminum are shown in [Table XII](#), and the price history is given in [Table XIII](#). The price of aluminum was very stable until the oil embargo of 1975. Until then, the rate of improvement of production efficiency overcame the effects of inflation. The largest market for aluminum comprises containers and packaging, with beer and beverage cans being the largest segment. Alloys 3004 and 5182 are used for can bodies and ends, respectively. The light weight and the high

TABLE XII Aluminum Consumption in the United States (Million Metric Tons)

Market	1965	1970	1975	1980	1985	1990	1995	1998
Containers and packaging	0.30	0.67	0.91	1.51	1.86	2.16	2.31	2.27
Building and Construction	0.87	1.02	1.01	1.19	1.38	1.21	1.22	1.39
Transportation	0.81	0.70	0.78	1.02	1.36	1.45	2.64	3.25
Electrical	0.49	0.62	0.55	0.64	0.64	0.60	0.66	0.72
Consumer durables	0.39	0.42	0.34	0.40	0.48	0.51	0.62	0.73
Machinery and Equipment	0.26	0.27	0.29	0.38	0.38	0.45	0.57	0.63
Other	0.31	0.36	0.25	0.27	0.28	0.26	0.28	0.27

From "Aluminum Statistical Review, 1965–1998" Aluminum Association, Washington, D.C.

recycle rate make aluminum cans very energy efficient and environmentally attractive containers. Foil, plastic, and paper laminants and pouches are used for packaging a wide variety of food and nonfood products.

The second largest market comprises building and construction. The 3000 series alloys are employed for residential siding, gutters, downspouts, and industrial and farm roofing and siding. Doors and windows are generally produced from 6063 extrusions. Screening, nails, and other fasteners are made from high-strength 5000 and 6000 series alloys. Anodized and colored aluminum panels are used in the curtain wall structures of many buildings and store-fronts. Aluminum's good resistance to weathering, high strength, and low weight for ease of handling account for its use in these applications.

The transportation field is another large market for aluminum. Heat-treatable alloys with high strength and toughness are used extensively in commercial and military aircrafts. The use of aluminum in automobiles to reduce weight results in increased fuel economy and improved performance. Aluminum is used in truck cabs and bodies, for wheels, and in some engine components to reduce weight and increase payload. Aluminum is used in rapid-transit vehicles, railway freight and tank cars, diesel locomotives, and buses to increase payload and save fuel. Aluminum is the preferred material for highway signs, divider rails, and highway lighting standards because of low maintenance and less damage to vehicles striking them. Aluminum-magnesium and aluminum-magnesium silicide alloys are widely used in marine service. Aluminum is also used in space vehicles because of its strength, toughness, lightness, weldability, and reflectivity.

Aluminum is used almost exclusively for high voltage electrical transmission lines. Aluminum has a volume electrical conductivity 62% of that of copper, but because of its lower density an aluminum conductor weighs only half as much as a copper conductor of equal current-carrying capacity. Aluminum is used in the rotor of many induction motors, for transformer windings, for bus

bars, and for the electrical conductor in most integrated circuits.

Aluminum has many applications in the chemical and petrochemical industries, where it is chosen for its resistance to chemicals, its low grasses such as *Distans stricta* are particularly successful. The residue contains numerous potentially valuable ingredients such as soda, iron oxide, and titania, but to date all attempts to extract these materials economically have been frustrated.

Aluminum smelting cells generate CO and CO₂ during normal operation but generate CO, CF₄ and C₂F₆ during anode effect. Although the amounts emitted are relatively small, these gases all contribute to global warming. The worse offenders are CF₄ and C₂F₆, which for equal amounts, produce 6500 and 9200 times the global warming compared to CO₂. There is an ongoing effort in the aluminum industry to minimize the frequency and duration of anode effects and thereby minimize the production of CF₄ and C₂F₆.

Aluminum smelting cells generate also gaseous and particulate fluorides. Cells are equipped with hoods to collect the evolved gases, which are then ducted to pollution control equipment. The early electrostatic precipitators and wet scrubbers have been largely replaced by dry scrubbers that trap the fluorides on alumina, which is then returned to the cells.

The cathode linings of aluminum cells must be replaced periodically. These spent linings represent the largest volume of solid waste associated with the smelting process. Because they contain fluorides and cyanide, they must be stored either under roof until they can be disposed of in an environmentally acceptable manner. The current method is to grind the linings, mix with lime, and incinerate. This captures the fluoride and destroys the cyanide, but is costly and increases the bulk volume that must be land filled. Current research is directed toward destroying the cyanide and recovering other valuable components from the spent lining. Some potential uses are as a flux in the steel industry, in making rock wool, supplemental

TABLE XIII Price History of Aluminum (Average Annual Price of Ingot)^a

Year	\$/kg Al	1999\$/kg Al
1950	.39	2.64
1955	.52	3.17
1960	.57	3.16
1965	.54	2.79
1970	.63	2.66
1975	.87	2.64
1980	1.59	3.14
1985	1.05	1.59
1990	1.65	2.06
1991	1.31	1.57
1992	1.18	1.37
1993	1.18	1.33
1994	1.57	1.73
1995	1.89	2.02
1996	1.57	1.64
1997	1.70	1.73
1998	1.44	1.44
1999	1.45	1.42
2000		

^a Prices to 1970 are from the *Historical Statistics of the U.S., Colonial Times to 1970*, U.S. Department of Commerce (1975). Prices from 1976 to 1999 are from the *U.S. Bureau of Mines, Mineral Commodity Summaries* (1980 and 1999). The annual price indices used to correct for inflation were supplied by the American Institute for Economic Research, Great Barrington, MA.

fuel for cement manufacture, and fuel in fluidized-bed boilers.

XIV. ENVIRONMENTAL CONSIDERATIONS

The production of aluminum creates several potential environmental problems. One is the disposal of bauxite residue. This residue is normally discharged into a drained, clay and polymeric membrane-lined waste reservoir. Eventually the reservoir becomes full. It can be restored to an acceptable grass field or even to agricultural production by capping with soil mixed with waste paper mill pulp and/or chopped hay and manure. Salt-tolerant grasses such as *Distans striata* are particularly successful. The residue contains numerous potentially valuable ingredients such as soda, iron oxide, and titania, but to date all attempts to extract these materials economically have been unsuccessful.

Aluminum smelting cells generate gaseous and particulate fluorides. Cells are equipped with hoods to collect the evolved gases and duct them to pollution control equipment. The early electrostatic precipitators and wet scrubbers have been largely replaced by dry scrubbers that trap the fluorides on alumina, which is then returned to the cells.

When the concentration of alumina in the electrolyte of the cell falls too low, a phenomena called anode effect occurs. During anode effect the cell emits CF₄ and C₂F₆. Both gases are potent greenhouse gases. The industry is working vigorously to reduce the frequency and duration of anode effects in order to minimize production of these gases.

The cathode linings of aluminum cells must be replaced periodically. Because these linings contain fluorides and cyanide, they must be kept dry until disposal. The conventional method of disposal is incineration in the presence of excess lime. The large volumes involved make this procedure costly. Research is being directed toward destroying the cyanide and recovering other valuable components from the spent lining. Some potential uses are as a flux in the steel industry, in making rock wool, supplemental fuel for cement manufacture, and fuel in fluidized-bed boilers.

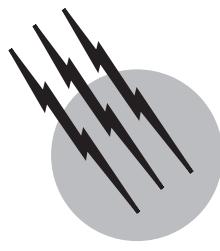
SEE ALSO THE FOLLOWING ARTICLES

ELECTROCHEMICAL ENGINEERING • ELECTROCHEMISTRY

- METAL FORMING • METALLURGY, MECHANICAL
- MINERALOGY AND INSTRUMENTATION • MINING ENGINEERING • TIN AND TIN ALLOYS

BIBLIOGRAPHY

- "Aluminum Standards and Data (1989)." The Aluminum Association, Inc., Washington, D.C.
- Berk, R., Lax, H., Prask, W., and Scott, J. (1982). In "Aluminum: Profile of the Industry" (J. Keefe, ed.) McGraw-Hill, New York.
- Burkin, A. R., ed. (1987). "Production of Aluminum and Alumina" Wiley, New York.
- Grjotheim, K., and Kuande, H. (1986). "Understanding the Hall-Héroult Process for Production of Aluminium," Aluminium-Verlag, Dusseldorf, Germany.
- Grjotheim, K., Krohn, C., Malinovsky, M., Matiasovsky, K., and Thonstad, J. (1982). "Aluminum Electrolysis: Fundamentals of the Hall-Héroult Process," 2nd ed. *** Aluminium-Verlag GmbH, Dusseldorf, GFR.
- Hatch, J. E., ed. (1984). "Aluminum, Properties and Physical Metallurgy," American Society for Metals, Metals Park, Ohio.



Carbon Fibers

J. B. Donnet

Ecole Nationale Supérieure de Chimie de Mulhouse

O. P. Bahl

National Physical Laboratory

Roop C. Bansal

Panjab University

T. K. Wang

Ecole Nationale Supérieure de Chimie de Mulhouse

- I. Classification of Carbon Fibers
- II. Raw Material for Carbon Fibers
- III. Carbon Fibers from PAN
- IV. Carbon Fibers from Pitch
- V. Carbon Fibers from Cellulose
- VI. Vapor-Grown Carbon Fibers (VGCF)

- VII. Carbon Fiber Structure
- VIII. Surface Treatment of Carbon Fibers
- IX. Properties of Carbon Fibers
- X. Carbon Fiber Demand-Pattern
- XI. Further Agenda
- XII. Applications of Carbon Fibers

GLOSSARY

- Activated carbon fibers** Carbon fibers that have been suitably processed to increase their adsorption.
- Carbon-carbon composites** All carbon composites, consisting of carbon fibers and carbonaceous matrix.
- Intermolecular reaction** Cross-linking reactions between two adjacent molecular chains.
- Intramolecular reaction** Reactions that occur within a molecule itself.
- Macropores** Pores with diameters greater than about 50 nm.

- Micropores** Pores with diameters less than about 2 nm.
- Preferred orientation** Average angle that the molecular chains make with fiber axis.
- Stabilization of polyacrylonitrile** Low temperature heating of polyacrylonitrile (PAN) in the presence of air, resulting in a ladder polymer structure.

CARBON FIBERS are very thin filaments, having diameters between 5 and 10 μm and consisting of 99.9% chemically pure carbon. They are not chemical compounds or alloys like many conventional structural materials such

as aluminium or steel, which are invariably associated with small quantities of other alloying elements. Carbon fibers were discovered nearly a century ago, when Thomas Edison first used them as filaments in incandescent lamps. These early carbon fibers were prepared by the carbonization of bamboo and rayon. Soon after the development of tungsten filaments, this use of carbon fiber was abandoned and they were almost forgotten.

The present breed of high performance carbon fibers was first prepared from rayon at the Wright Patterson Air Force Base for the aerospace industry to provide lightweight and more fuel-efficient aircraft capable of high speeds. Union Carbide developed the process and commercialized the product as Thorne. At about the same time, Shindo in Japan and Watt and coworkers in the United Kingdom, working independently, successfully prepared carbon fibers from polyacrylonitrile (PAN). The pitch-based carbon fibers were developed by Otani, who used polyvinylchloride (PVC) pitch as a raw material, while mesophase pitch fibers were introduced by Singer at Union Carbide in the United States.

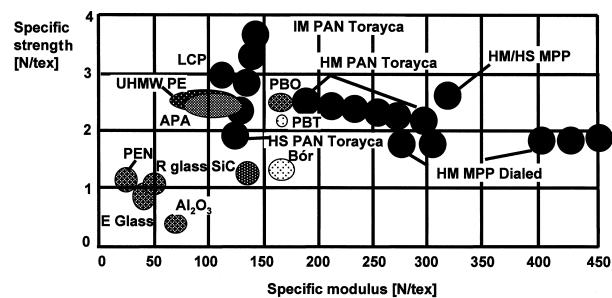
Carbon fibers exhibit both a high specific strength and a high specific modulus. Moreover, carbon fibers are available in a wide range of strength (between 2000 and 7000 MN/m²) and modulus (between 200 and 900 GN/m²). Polyacrylonitrile-based carbon fibers form the most important group, with high strength and high modulus, while the mesophase-pitch-based fibers provide an ultra-high modulus type although with lower strengths. Isotropic pitch fibers have both a low modulus and low strength and can be used as general purpose carbon fibers.

In addition to strength and stiffness, carbon fibers are elastic to failure at normal temperature, which makes them creep-resistant and susceptible to fatigue. They have exceptional thermophysical and excellent damping characteristics. They are electrically conductive, and their conductivity can be enhanced by intercalation. They are chemically inert except in strongly oxidizing environments or when in contact with certain molten metals.

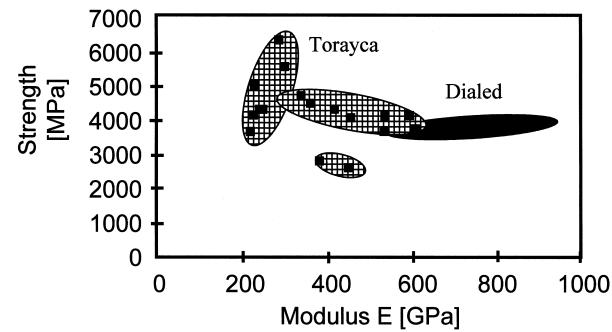
I. CLASSIFICATION OF CARBON FIBERvhS

After considerable discussion and debate, the International Union of Pure and Applied Chemistry (IUPAC) International Committee has classified carbon fibers into five types on the basis of scientifically verifiable theoretical values:

1. Ultra-high modulus (UHM) type: Carbon fibers with modulus greater than 500 GPa;
2. High modulus (HM) type: Carbon fibers with modulus greater than 300 GPa and with a tensile strength to tensile modulus ratio of less than 1%;



(a)



(b)

FIGURE 1 Mechanical properties of some commercial carbon fibers: (a) specific strength versus modulus and comparison with other materials; (b) strength versus modulus for Torayca™ (PAN-based, Toray Industries) and Dialed™ (mesophase pitch-based, Mitsubishi Chemical Co.) carbon fibers. [From: Korinek, Z. (1999). *Acta Montana IRSM AS CR, Series B*, **9**, 19–23.]

3. Intermediate modulus (IM) type: Carbon fibers with modulus up to 300 GPa and with strength-to-modulus ratio greater than 1×10^{-2} ;
4. Low modulus type: Carbon fibers with modulus as low as 100 GPa and with low strength; these are fibers with isotropic structure;
5. High tensile (HT) type: Carbon fibers with tensile strength greater than 3000 MPa and strength-to-modulus ratio between 1.5 and 2×10^{-2} .

Fig. 1 shows the strength and the modulus positions of some commercial carbon fibers expressed either in specific (Fig. 1a) or in absolute (Fig. 1b) values.

II. RAW MATERIAL FOR CARBON FIBERS

Carbon fibers can be prepared using any fibrous precursor with a carbon backbone that can yield carbonaceous residue on pyrolysis. The mechanical properties of carbon fibers, however, are determined by the nature of the precursor used and the processing conditions employed. A precursor should preferably possess the following characteristics:

TABLE I Weight Loss on Heating at 1000°C in Helium for Different Precursor Fibers

Polymer type	Weight loss total (%)
Rayon	87
Polyacrylonitrile	60
Polyacrylonitrile, preoxidized	38
Saran	74
Poly(vinyl alcohol)	93
Ramie	91
Pitch	30

- Easily spun into filamentary form;
- Able to decompose without melting and at a slow rate;
- High crystallinity with a high degree of preferred orientation along the fiber axis;
- Produce minimum loss of carbon as volatiles during pyrolysis and maximum carbon yield.

The significant factors in obtaining a high carbon fiber yield are the nature and the mechanism of the pyrolysis process, the capacity of the precursor for cyclization, ring fusion, coalescence, and the nature of the stabilizing pre-treatment. The results of the thermogravimetric analysis carried out on several polymeric fibers at 1000°C in helium atmosphere (Table I) indicate that the yield of the carbon fiber is maximum in the case of PAN and pitch precursor fibers.

Several organic polymeric materials such as PAN, rayon, aromatic heterocyclic polymers, linear polymers, and pitch satisfy some of these requirements and, consequently, have been investigated as starting materials for the preparation of carbon fibers. For large-scale production, however, the choice has been narrowed down to three main types: PAN-based, pitch-based, and rayon-based. At present, the PAN-based carbon fibers constitute the most important group and are highly desirable in high performance composites for aircraft, aerospace, and other highly technical applications. The pitch-based carbon fibers possess unique structural characteristics that have opened up a variety of new applications. New pitch materials producing better carbon fibers and new processes for obtaining better materials from existing pitch sources have been developed. In the case of rayon-based carbon fiber, the yield is very low (10–15%; see Table I), and the production of high-performance materials requires a difficult and costly high-temperature stress treatment. Thus their production is quite limited.

The details of the processes used for the manufacture of carbon fibers from different precursors are not the same, although their essential features are similar. They involve a stabilizing treatment to prevent melting or fusion of the

precursor fiber, a carbonizing heat treatment to eliminate the noncarbon elements, and a high temperature graphitizing treatment to enhance the mechanical properties of the final carbon fibers. In order to produce carbon fibers with a high degree of preferred orientation along the fiber axis (which determines their mechanical properties), the precursor fibers are given a stretching treatment during one of the three stages of the pyrolysis process (Fig. 2). This orientation is generally carried out during the oxidative stabilization step in PAN fibers and during the graphitization stage in the case of rayon fibers.

Another type of carbon fiber which is receiving increasing interest is vapor-grown carbon fiber. These carbon fibers can have exceptional thermal and electrical properties.

III. CARBON FIBERS FROM PAN

Polyacrylonitrile fibers are the most widely used precursor for the manufacture of present-day carbon fibers. The advantages are a high degree of molecular orientation, higher melting point, and a higher yield of carbon. The fiber can be spun under stretch, and the degree of stretch can be as high as 800%. Polyacrylonitrile fibers give rise to a thermally stable, highly oriented molecular structure when subjected to a low-temperature heat treatment, and this structure is not significantly disrupted during the carbonization treatment at higher temperatures.

Polyacrylonitrile has highly polar nitrile groups which cause strong dipole-dipole forces that act as cross-links, making the polymer soluble only in highly ionizing solvents, increasing its melting point, and making it more suitable as a carbon fiber precursor. These dipole-dipole interactions, however, prevent the rearrangements necessary for the development of crystallinity and result in shrinkage of the fiber on heat treatment. Because this shrinkage disturbs the orientation in the fiber, it must be avoided. Thus, special quality PAN fibers are prepared which should have the following minimum characteristics:

- Denier: 1.0–1.5
- Co-monomer content: 2–3%
- Molecular weight: $M_n \sim 2.5 \times 10^4$, $M_w \sim 25 \times 10^4$
- Crystalline content: >55%
- No surface finish

A. Oxidative Stabilization of PAN Fibers

Oxidative stabilization of PAN precursor fiber is an important and most time-consuming step in the preparation of the carbon fibers. Improper stabilization results in

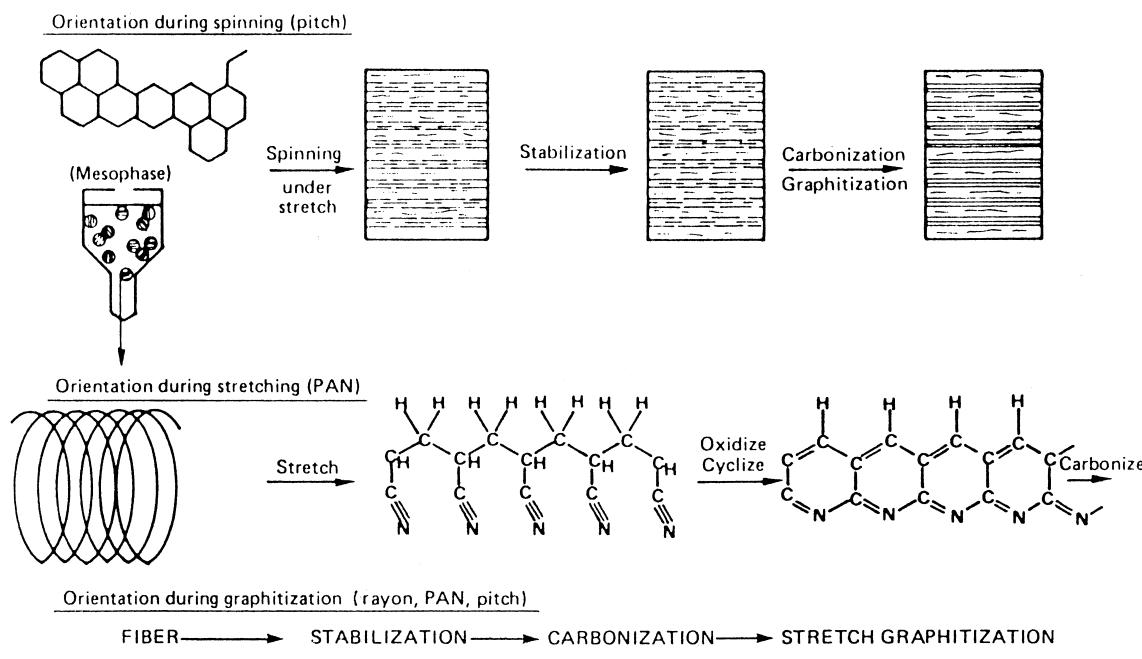


FIGURE 2 Methods for creating preferred orientation in carbon fibers.

incomplete oxidation and may cause a “blowout” of the core portion of the fiber during carbonization. When carried out in air or in the presence of oxidizing agents between 220 and 250°C in a batch process or by passing the precursor tow continuously through several heating zones in a furnace, the stabilization takes several hours. To reduce the time of stabilization, several modifications have been suggested to the conventional stabilization procedures. It has been observed that the stabilization rate is enhanced by modifying the chemical composition of the precursor fiber, either by the addition of a comonomer or an additive or by impregnation with certain chemicals. Polyacrylnitrile precursor fibers require only 25 min for stabilization when spun from a PAN copolymer containing about 2% methacrylic acid and when a V-shaped temperature profile is used. Similarly, a time period between 2 and 20 min has been claimed for stabilizing by contacting the precursor fiber with a heated surface in the presence of an oxidizing agent, making it possible to couple spinning and stabilization.

The PAN copolymer used for preparing precursor fibers usually contains between 2 and 10% of a comonomer, which could be methylacrylate (MA) or methyl methacrylate (MMA), along with a small amount of a third comonomer which initiates the stabilization reactions at lower temperatures. The comonomer MA acts as a plasticizer to aid the spinning process and facilitates subsequent stretching of the fiber during the oxidation treatment. The fibers are stretched about 2.5 times in a coagulation bath and are further stretched by about 14 times in steam at

100°C after washing. This stretch aligns the molecular chains along the fiber axis and enhances the Young's modulus of the final carbon fibers.

During stabilization under heat treatment in a controlled oxidizing atmosphere at about 220°C under tension, several thermally activated processes take place, causing a considerable reorganization of the polymer chains and three-dimensional linking of the parallel molecular chains by oxygen bonds. These oxygen cross-links keep the chains straight and oriented parallel to the fiber axis even after the release of tension after the treatment.

B. Shrinkage during Stabilization

The oxidative stabilization of PAN fibers at around 220°C in air under a constant load results in length changes. The fibers are found to shrink under low loads while they initially increase in length and then shrink under high loads. These length changes are attributed to molecular relaxation, cyclization, and dehydrogenation reactions. An overall shrinkage between 13 and 34% has been observed. According to Fitzer, the overall shrinkage is made up of a physical contribution called entropy shrinkage due to molecular relaxation and a chemical contribution called chemical reaction shrinkage due to cyclization and dehydrogenation reactions. A part of the chemical shrinkage may also be due to the cross-linking reactions caused by interaction with the oxygen or the comonomer. The two contributions can be clearly distinguished by a minimum in the first derivation curve (Fig. 3). While the entropy shrinkage

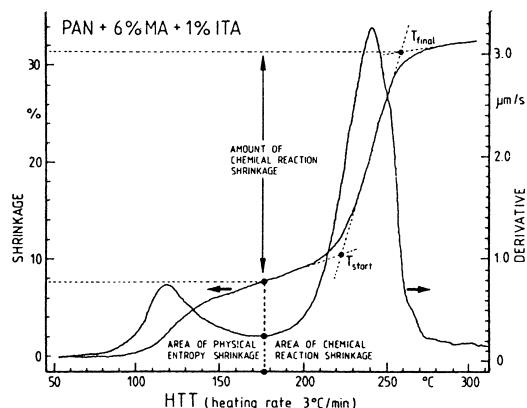


FIGURE 3 Typical shrinkage curves measured during stabilization of PAN fibers showing physical entropy shrinkage and chemical reaction shrinkage contributions. [From Fitzer, E., Frohs, W., and Heine, M. (1986). *Carbon* **24**, 387.]

is completed below 200°C , the chemical shrinkage starts thereafter. The latter is influenced by the presence of the copolymer and by an increase in the heating rate. At higher heating rate, the chemical shrinkage starts later but takes place at an enhanced rate. This is indicative of the cross-linking reaction's contribution to the basic shrinkage caused by cyclization and dehydrogenation reactions.

The molecular relaxation and, hence, the entropy shrinkage can be minimized either by constraining the precursor longitudinally to constant length or by restraining under constant force during the stabilization heat treatment. In a continuous process, the stabilization can be carried out by feeding and removing the fiber at the same speed from the heat-treatment zone so that the fibers can undergo compensating local shrinkage and extension during the treatment.

C. Chemistry of Oxidative Stabilization

Several techniques, such as infrared spectroscopy, thermal analysis (DSC), gas chromatography, and mass spectrometry have been used to understand the chemistry involved in the oxidative stabilization process. However, the exact nature of the reactions is still not clear, because the linear PAN molecule contains unsaturated side groups which cause additional polymerization reactions due to the presence of such additives as comonomer, by-products and traces of initiators from the earlier polymerization, end groups and ionic groups introduced during the spinning of the fiber, the rate of heating, and the atmosphere in which the stabilization is carried out. It is, however, clear that the oxidative stabilization involves cyclization, dehydrogenation, and oxidation, which transform the thermoplastic chain molecule into a thermally stable cyclized

structure often referred to as a ladder polymer. Several different structures have been proposed for this cyclized ladder polymer, but the view that the pendant nitrile groups in PAN cyclize to give polyimine structures appears to be generally accepted. The evolution of major volatiles such as H_2O , CO_2 , and HCN and the formation of the stabilized polymer during the stabilization process is shown in Fig. 4. When properly stabilized, the stabilized polymer should contain between 10 and 15% oxygen. This oxygen is present as $\text{C}-\text{OH}$ and $\text{C}=\text{O}$ groups and promotes cross-linkings and helps in the aromatization of the cyclized sequences by the elimination of water as a result of the reaction of oxygen with the backbone $-\text{CH}_2$ groups. The presence of these groups has been established by infrared spectroscopic studies (Fig. 5) of such samples. An oxygen content greater than 12% causes deterioration of the fiber quality, while an oxygen content less than 8% results in a low carbon fiber yield due to excessive weight loss during the subsequent carbonization process.

Another important parameter is the temperature at which the stabilization should be carried out. Usually, temperatures between 200 and 300°C have been mentioned. When the stabilization is carried out isothermally at lower temperatures, the reaction is very slow and needs longer times, while at higher temperature, the heat evolution during stabilization is so large that it can cause severe damage to the fiber. On the basis of a comprehensive study involving a large number of experiments, Fitzer has suggested that final temperature of 270°C , with a heating rate of 5°C min^{-1} up to a temperature of 230°C , where the chemical reactions are said to start, and a heating rate of 1°C min^{-1} thereafter up to 270°C to avoid overheating of the fiber are the optimum conditions for the best oxidative stabilization of PAN precursor fibers. Of course, these conditions will vary depending on the type of the precursor used.

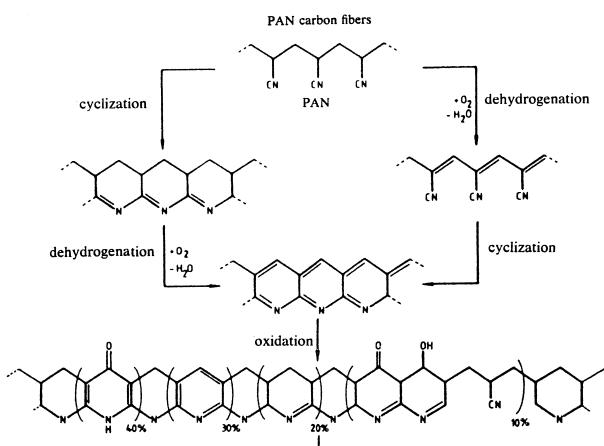


FIGURE 4 Chemistry of PAN stabilization. [From Fitzer, E., Frohs, W., and Heine, M. (1986). *Carbon* **24**, 387.]

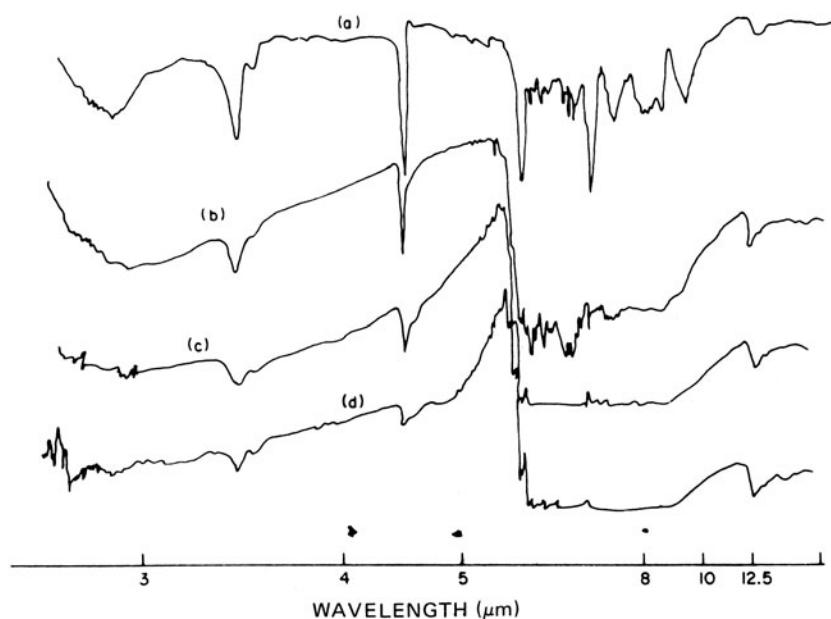


FIGURE 5 Infrared spectra of PAN fibers (copolymer) heat-treated in oxygen at 215°C for different durations (a) 0 min, (b) 25 min, (c) 200 min, and (d) 400 min. [From Manocha, L. M., and Bahl, O. P. (1980). *Fibre Sci. Technol.* **13**, 199.]

D. Pyrolysis of Pan Fibers

The stabilized PAN fibers are subjected to thermal pyrolysis in an inert atmosphere when the noncarbon elements are eliminated as volatiles with a carbon fiber yield of about 50%. The pyrolysis can be divided into three stages involving different chemical reactions. These three stages are the three temperature ranges 300–600, 600–1300, and 1300–2500°C.

The first heating stage (temperature range 300–600°C) is critical as it involves chemical reactions leading to the evolution of H₂O, NH₃, HCN, and CO₂, together with hydrocarbons in the form of tar. The release of tar depends upon the degree to which the fiber structure has been stabilized. If the structure is not well stabilized, the release of larger quantities of tar results in a low carbon fiber yield. The evolution of volatiles depends on the rate of heating, which has to be kept low ($1.5^{\circ}\text{C min}^{-1}$). A faster release of volatiles damages the fiber. During this stage, the hydroxyl groups present in the oxidized PAN fiber initiate cross-linking condensation reactions, which help in the reorganization and coalescence of the cyclized reactions (Fig. 6). This cross-linking fixes the structure of the polymer, while the remaining linear segments either become cyclized or undergo chain scission reactions evolving the gaseous volatiles.

In the second stage of pyrolysis (600–1300°C) the rate of heating can be enhanced as the major reaction products have already been removed. During this stage, the cyclized structure undergoes dehydrogenation reaction and starts linking in the lateral direction bounded by nitrogen

atoms. These focused, oriented heterocyclic molecules coalesce into graphite-like ribbons and fibrils with the evolution of nitrogen and hydrogen (Fig. 6). The elimination of nitrogen continues above 700°C and reaches a maximum rate at 900°C. At 1300°C the carbon fiber is left with a very low nitrogen content (2–3%). The carbon fiber yield, which is about 50% after the second stage of pyrolysis, can be enhanced by carbonizing in an atmosphere of HCl vapors. Hydrochloric acid inhibits the formation of HCN by removing nitrogen as ammonia and also has a marked dehydrating effect. A small degree of shrinkage of the fiber has also been observed during carbonization, even when the precursor fiber is fully stabilized.

The third stage of pyrolysis (1300–2500°C) is also known as graphitization because the structure of the carbon fiber tends to become more and more graphite like. Heat treatment in this step, which involves a few seconds only even beyond 2500°C, does not result in any appreciable loss in weight but improves the orientation of the crystallites in the direction of the fiber axis. This enhances the Young's modulus of the carbon fiber. The Young's modulus of the carbon fiber has been found to be directly related to the final heat treatment temperature during graphitization.

E. Post Spinning Modification

It has been established by Bahl and coworkers that there is a direct relationship between the quality (particularly primary Young's modulus) of the precursor and the

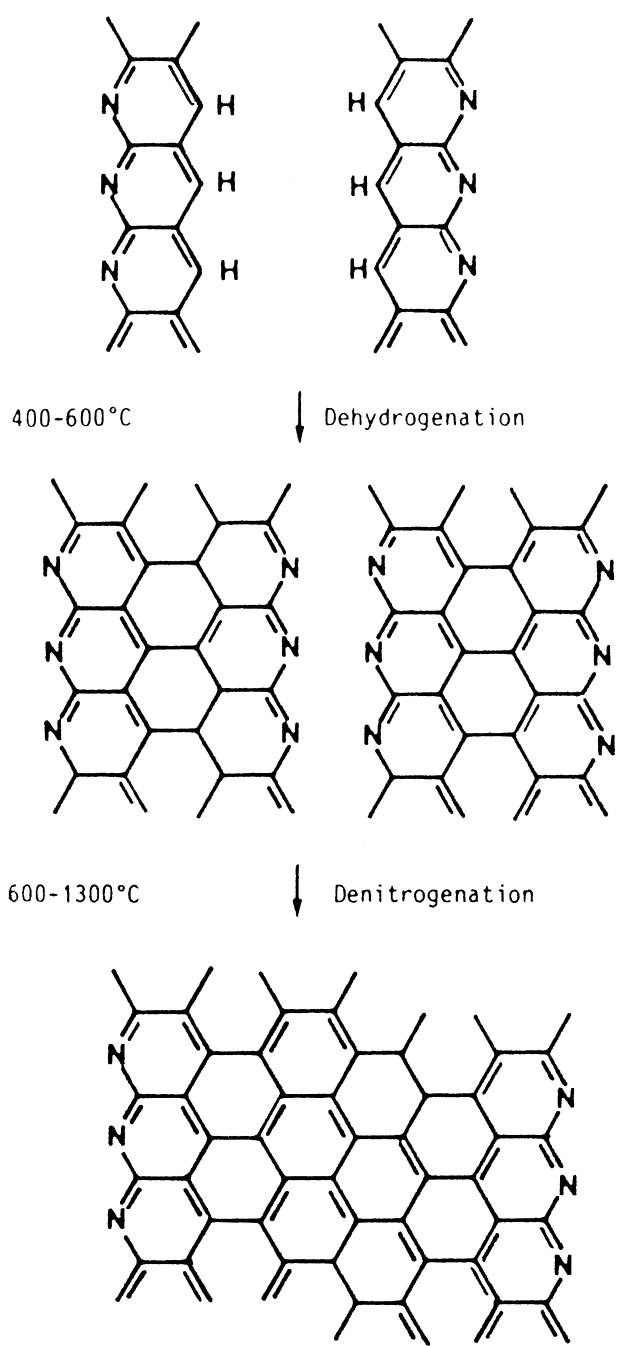


FIGURE 6 Dehydrogenation and denitrogenation process during the pyrolysis of PAN [From Donnet, J. B., and Bansal, R. C. (1988). "Carbon Fibers," Marcel Dekker, New York.]

mechanical properties of the carbon fibers derived therefrom. Also there is a limit on the ultimate stretching that can be imparted during the spinning of the precursor. In order, therefore, to overcome these limitations, post spinning modifications on PAN fibers have been carried out which have resulted in improvement in the qual-

ity of PAN precursor vis-à-vis the resulting carbon fiber properties.

These modifications can be put into three broad categories, namely

1. Modification through coatings with certain resins to suppress exotherm during the cyclization reaction;
2. Modification using certain reagents like Lewis acids, organic bases, etc., which act as catalysts for the cyclization reaction;
3. Post spinning stretching using different plasticizers like DMF, CuCl, nitrogen, etc., to improve the structure of PAN.

The above modifications can help in

- decreasing the activation energy of cyclization;
- reducing the stabilization exotherm;
- increasing the speed of cyclization reaction;
- improving the orientation of molecular chains in the precursor fiber;
- reducing sheath–core formation in stabilized PAN fibers.

Post spinning modifications not only modify the characteristics of the PAN fibers but also increase the mechanical properties of resulting carbon fiber as well. There is still a large gap between tensile strength achieved so far in commercially available PAN-based carbon fibers (~7.3 GPa) and the theoretically expected value (~100 GPa). For bridging this gap, post spinning modifications, in addition to control of processing parameters and precursor quality, etc., are expected to play a key role in realizing the ultimate goal.

IV. CARBON FIBERS FROM PITCH

Pitch is an important, abundantly available, and low-cost material for the production of carbon fibers. Its suitability for carbon fiber manufacture is determined by its ability to spin and be converted into a nonfusible state. These properties are largely dependent on its chemical composition and molecular weight distribution, which can be controlled by the addition of certain substances or by extraction with certain solvents or by distillation.

There are two types of pitch carbon fibers: isotropic-pitch-based and mesophase-pitch-based. Isotropic-pitch-based carbon fibers have poor mechanical properties and have no importance for advanced composites. They can be used as general purpose fibers for insulation, as fillers in plastics, and in strengthening concrete, etc.

A. Isotropic Pitch Carbon Fibers

The first pitch-based carbon fibers were prepared by Otani by the pyrolysis of PVC pitch. The molten pitch was extruded through a pinhole, and the filaments were extended and solidified while flowing down through the air and collected over bobbins at high speed.

The carbonization of pitch fibers involves, as usual, an oxidation pretreatment in ozonized air below 60°C and a subsequent heat treatment in air at the rate of 1.5°C min⁻¹ up to 60°C, keeping the final temperature for about 3 hr. The ozone treatment results in the addition of ozone to the aromatic nucleous, producing ozonide- or peroxide-type linkages which create a kind of intermolecular cross-linking. In addition, some ring scission reactions also take place. These cross-linking reactions render the filament infusible during further heat treatment. Several other oxidizing agents such as air, oxygen, and nitrogen dioxide have been used, but the rate of oxidation is found to be slow and higher temperature is required.

Thermally stable fibers obtained after the preoxidation treatment are carbonized in a stream of nitrogen by heating at the rate of 5°C min⁻¹ up to final temperature varying between 1000 and 1300°C, which is maintained for about 15 min. The total weight loss during preoxidation and carbonization was 24% of the original pitch filament.

B. Mesophase Pitch Carbon Fibers

The most important stage in the process is the preparation of a mesophase pitch suitable for making high-performance carbon fibers. Initially prepared mesophase-pitch carbon fibers could not find large-scale applications primarily because of

- its prohibitive cost;
- the availability of limited numbers of suitable mesophase pitches;
- the strength of these carbon fibers was insufficient whereas Young's modulus could be increased at will and this resulted in strain to failure of less than 0.3%.

The preparation of such mesophase pitches is a part of the patented literature. The essentials of several methods are given below, although the details are not clear.

The method patented by Union Carbide involves slow heating of an isotropic petroleum pitch in nitrogen atmosphere to a temperature between 360 and 400°C for 12–24 hr. During the heat treatment, under mild agitation, the light hydrocarbons and polycondensation products are vacuum- or nitrogen-stripped so that the higher molecular weight aromatics can condense to form a mesophase. The mesophase yield is between 45 and 70% and can be spun into fibers between 350 and 400°C. If, however, the pitch is first de-ashed and distilled to remove low-temperature

volatile matter and heated to get the mesophase, the material obtained is less viscous and can be spun more smoothly.

Exxon Research and Engineering patented a process in which the heavy molecular weight components are first removed by solvent extraction. The insoluble portion, which is called a "neomesophase former," softens between 230 and 400°C, producing mesophase. This mesophase called the neomesophase, has a lower softening point than the conventional mesophase pitch and can thus be spun at a lower temperature. The amount of neomesophase former is enhanced by vacuum-stripping of the pitch at 420°C before extraction. In a similar process developed by Great Lakes Carbon, the coal tar pitch is extracted with creosote oil and then vacuum-stripped at 130°C to eliminate low boiling constituents. The residue produces mesophase when heat-soaked between 280 and 350°C for 20 hr.

As a result of sustained R & D efforts by Mochida and Mitsubishi Gas Chemical Company, a relatively simple method of preparing mesophase pitch from pure aromatic hydrocarbons has been invented in which HF/BF₃ is used as a condensation catalyst. The catalyst is fully recovered for repeated use. The preparation conditions in terms of temperature and time, as well as starting monomers control the structure and properties of the mesophase.

Main features of the preparation are

- Nondehydrogenative condensation of aromatic nuclei (parent monomer), producing naphthenic groups by every reactor condensation are thus dehydrogenated to an extent which depends on time and temperature of preparation. The big advantage is that the mesophase pitch so prepared is extraordinarily rich in naphthenic groups.
- The final mesophase pitch reflects the structure of the starting monomer and hence a variety of properties are controlled by the starting monomer.
- Yield of the mesophase can be up to 90% and the pitch is extremely homogeneous in nature

Mesophase pitch prepared by the catalytic method can be modified further both chemically and physically. Molecular structure of mesogens is changed by addition, substitution, and oxidation. Since the mesophase pitch possesses a high degree of solubility, modification reaction takes place in the liquid homogeneous phase.

C. Spinning of Mesophase Pitch

Spinning of mesophase pitch is a difficult task and poses several problems since

- the mesophase is always in transition and has a tendency to be transformed into coke;

TABLE II Pitch Spinning Parameters Used by Two Companies

	Kureha	Union Carbide
Spinning temperature(°C)	240–230	350–436
Spinneret diameter (nm)	0.1–0.9	0.7–0.38
Number of holes	30–120	41–1,000
Spinneret capillary length-to-diameter ratio	—	2–5
Filament diameter (μm)	8–30	8–50
Filament speed (m/min)	400–1,700	29–226
Draw ratio	400–5,000	3–1,702

- it is too viscous for continuous smooth spinning;
- as the spinning temperature is generally higher compared to an isotropic pitch, there is evolution of gas which may be entrapped in the fiber and produce defects;
- mesophase pitch has a heterogeneous structure consisting of anisotropic regions.

Several spinning modes such as centrifugal spinning, jet spinning, and conventional melt spinning have been used to spin mesophase pitches. Continuous multifilament mesophase fibers are obtained using commercial melt spinning equipment. The spinning is carried out through a 0.3 mm diameter orifice, so that for spinning 10 μm diameter filaments at 2.5 msec⁻¹, the draw ratio is about 1000 : 1. The important parameters for spinning mesophase pitch fibers are summarized in **Table II**.

The as-spin mesophase pitch fibers are highly anisotropic and show elongated anisotropic domains when viewed under a polarized light microscope. X-ray diffraction patterns show that about 75% of the larger planes lie within an angle of $\pm 15^\circ$ of the fiber axis, indicating a high degree of preferential alignment.

D. Stabilization and Pyrolysis of Mesophase Fibers

Because mesophase pitches are thermoplastic, the precursor fibers need to be stabilized by a suitable oxidation treatment. This stabilization is generally carried out between 250 and 350°C in electrically heated air ovens for about 40 min.

The stabilized fibers are pyrolyzed by passing continuously through a series of heating zones until the final carbonization temperature, which may be between 1000 and 2000°C, is reached. This causes the elimination of remaining heteroatoms as volatiles and helps to develop a larger structure. Increasing the final heat treatment tem-

perature merely enhances the preferred orientation and improves the mechanical properties of the carbon fibers.

When the carbon fibers are heat-treated in the temperature range 2500–3000°C, the Young's modulus approaches a value of 880 GPa, close to the graphite single crystal value of 1000 GPa. Still, this graphitization step is optional and is carried out only to obtain ultra-high modulus carbon fibers.

V. CARBON FIBERS FROM CELLULOSE

Cellulose is a naturally occurring material. It decomposes before melting, which makes it a suitable precursor for the production of carbon fibers. Generally, cellulose in the form of rayon or regenerated cellulose is used for carbon fiber manufacture because the basic composition, the degree of polymerization, and the spinning conditions can be controlled.

The pyrolysis is generally carried out at low temperature and in a reactive atmosphere such as oxygen, air, chlorine, or HCl vapors or in the presence of flame retardants. These chemical compounds promote dehydration of cellulose and inhibit the formation of tar, resulting in a greater yield of carbon fiber. The pyrolysis involves four different stages and the major degradation occurring between 210 and 320°C, resulting in a weight loss varying between 75 and 90%, depending on the raw material and the rate of heating.

The first stage of pyrolysis, up to 150°C, involves the desorption of water amounting to about 12%. This is followed by two competing processes. The first involves random dehydrogenation of hydroxyl groups as structural water (150–240°C), producing dehydro cellulose, the main source of carbon at high temperatures. The second reaction involves the cleavage of C–O and C–C bonds (240–400°C), leading to the formation of levoglucosan, which at high temperatures produces unwanted tarry products. The condensation and dehydrogenation reactions break up each cellulose unit into a four-carbon atom residue, which polymerizes to give a carbon polymer, ultimately producing a graphite-like structure with an ultimate carbon yield of less than 20%. The formation of levoglucosan, which reduces the carbon fiber yield by tar formation can be decreased by carrying out pyrolysis in the presence of flame retardant or in a reactive atmosphere. The carbon fiber yield has been found to be enhanced to 30–35% when pyrolysis is carried out in the presence of HCl vapors.

The carbonization of cellulose fibers occurs in the temperature range 1000–1500°C in an inert atmosphere, usually for a few seconds. Further heat treatment to higher

temperatures under tension improves the mechanical properties. The carbonized fibers are subjected to graphitization under stress above 2500°C for a fraction of second when the Young's modulus improves further.

VI. VAPOR-GROWN CARBON FIBERS (VGCF)

Vapor-grown carbon fibers bridge the size gap between nanotubes and commercial carbon fibers. These can be grown readily with dimensions from less than 30 nm to greater than 30 μm. Such carbon fibers are grown in hydrocarbon atmosphere with the help of fairly dispersed solid catalysts such as iron or other transition metals at high temperatures (typically 1100°C). Carbon from the hydrocarbon gas is absorbed into the catalyst and then precipitates out integrating itself into a fine filament of carbon with diameter approximately equal to that of the catalyst particle. Out of the various methods that are available for producing VGCF, two are most common, namely the conventional substrate method and more modern floating catalyst method.

Using substrate-bound catalyst, very long and slender fibers are grown. Strong and graphitic filaments are obtained in slow reactions with alkanes, while filaments with less-ordered carbon are obtained from unsaturated hydrocarbons. These carbon filaments have a unique structure, consisting of an amorphous core surrounded by a graphitic skin and containing a metal component present as a particle in the tip of the filament or as a dispersion along the body of the filament. This form of VGCF, when subsequently heat-treated (Table III) exhibits a very high thermal conductivity of the order of 1800–1900 W/mK in addition to high modulus and high surface area. These fibers have been used to make high thermal conductivity composites using matrices such as aluminum, epoxy, carbon, and ceramics (SiC) for aerospace thermal management materials. Heat-treated VGCF in a carbon matrix at a 70%

fiber loading gives a thermal conductivity of 910 W/mK, which is more than double the value of copper and five times that of aluminum.

A. Floating-Catalyst Method

Obviously, catalyst particles are airborne in the reactant gas in this method. Very fine filaments are produced with a diameter between 0.1–0.2 μm and the length in the vicinity of 100 μm. Since the fibers are not held in place these usually get entangled during growth. A small addition of sulfur into the reaction has been found to increase the yield of VGCF substantially which in fact is a breakthrough. This happens since sulfur helps to activate the catalyst by forming a lower melting iron-sulfide eutectic which eventually allows easier diffusion of carbon through the catalyst particle.

Degree of graphitization is a crucial property from the point of view of making composites for thermal management applications. Table III gives degree of graphitization of the two types of VGCFs. Substrate-grown VGCFs show overall best graphitization which is totally dependent on the level of heat treatment given. On the other hand vapor-grown, floating-catalyst carbon fibers possess a fairly high degree of graphitization.

The floating-catalyst technique has provided submicron VGCFs, which in itself is a new class of carbon fibers providing potential for mechanical reinforcement combined with improved electrical properties. This has opened up new applications in rubber automotive panels, structural materials, aerospace components, thermal management, etc.

Keeping in view the simplicity of processing floating-catalyst VGCF as compared to obligatory well-controlled and multisteps processing in the case of PAN-based carbon fibers, it is projected that such VGCF will be much cheaper—about \$10/kg. Applied Sciences Inc. USA is underway to upscale their pilot plant (of 5 TPA) to about 200 TPA.

VII. CARBON FIBER STRUCTURE

The structure of carbon fibers has been widely studied using several different techniques such as wide- and small-angle X-ray and electron diffraction, transmission, scanning and surface replica electron microscopy, and optical microscopy. These techniques have provided useful information regarding the micro- as well as the macro-structural features of carbon fibers so that many of the aspects of carbon fiber structure are well understood, accepted, and agreed upon. There is general agreement that carbon fibers

TABLE III Changes in Degree of Graphitization with Heat Treatment

Heat treatment (°C)	Fiber type	D-spacing	Degree of graphitization (%)
2800	Substrate VGCF	0.33663	86
2500	-Do-	0.33770	73
2200	-Do-	0.34206	23
As-grown	-Do-	0.34490	—
As-grown	Floating catalyst VGCF	0.3384	64

- are composed of two-dimensional turbostratic graphitic crystallites oriented preferentially parallel to the fiber axis;
- have a two-phase structure with several kinds of micro-as well as macro-heterogeneity;
- from different precursor materials have many common structural features.

In addition, these techniques have provided qualitative and quantitative information about several structural parameters, such as crystallite size and orientation, stack height and width, the shape, size, and distribution of the pores, and the presence of three-dimensional order.

A. Crystallite Dimensions

Continued X-ray and electron diffraction studies suggest that carbon fibers heat-treated to high temperatures have a two-dimensional turbostratic structure with only a small portion of highly oriented crystallites so that there is no regular structure in the C direction. Thus, the crystallites have two size parameters: the crystallite width, La, and the crystallite height, Lc. These parameters depend largely on the nature of the precursor material, the heat-treatment temperature (HTT), and the amount of the strain during graphitization (Table IV). Both La and Lc increase with increase in the HTT, although the increase is smaller with HTTs below 1400°C. The crystallite di-

TABLE IV Crystallite Dimensions of Carbon Fibers

HTT (°C)	Lc (nm)	La (nm)
PAN-based		
1000	1.0	2.0
1400	1.8	3.5
2000	3.4	5.4
2400	4.0	6.2
2800	6.0	7.0
2250 ^a	20.0	50.0
Pitch-based		
1000	1.10	3.92
1500	1.52	7.13
2000	2.52	7.45
2500	2.69	6.85
2850	3.33	6.73
2850 ^b (40%)	4.40	8.00
2850 ^b (80%)	5.10	7.40
2850 ^b (120%)	8.90	13.60
2850 ^b (180%)	14.50	—

^a Strain-graphitized.

^b Strain-graphitized (% strain).

TABLE V Crystallite Width Parallel (La //) and Perpendicular (La ⊥) to the fiber Axis

HTT (°C)	Crystallite width (nm)	
	La //	La ⊥
PAN-based carbon fibers		
1250	3.5	3.4
2700	11.9	6.4
Rayon-based carbon fibers		
2800	5.9	5.0
2800 ^a	8.1	5.4
2900	13.0	6.5

^a Stress-graphitized.

mensions of stress-graphitized PAN carbon fibers can be as large as Lc 20 nm and La 50 nm. The axial crystallite width, La parallel (La //) and the transverse, La perpendicular (La ⊥) to the fiber axis from X-ray scattering measurements for several PAN and rayon-based carbon fibers (Table V) indicate that the layers are longer in the axial direction than in the transverse direction. The difference increases with increase in the HTT or the degree of preferred orientation. The rayon-based carbon fibers prepared at the same HTT but with varying preferred orientation show a constant transverse crystallite width but increasing axial width. This is attributed to the straightening of wrinkled ribbons of graphite layers under the influence of increased thermal versus mechanical treatment.

B. Crystallite Orientation in Carbon Fibers

Carbon fibers have a high degree of preferred orientation of the crystallites along the fiber axis. This structural parameter, in fact, is responsible for their excellent mechanical properties. Several preferred orientation parameters have been recognized. W. Ruland and A. Fourdeux measured the preferred orientation of a number of carbon fibers of different origins in terms of the orientation parameter *q* (which can be obtained from the measured orientation distribution function $1002(\phi)$ in wide angle X-ray diffraction) and observed that this orientation parameter varies between -1 for perfect orientation and 0 for random orientation of the crystallites. A.A. Bright and L.S. Singer measured preferred orientation as a function of HTT for two mesophase-pitch-based fibers in terms of full width at half-maximum (FWHM) and found that it decreased from 30 to 5° on heat-treatment at 3000°C, indicating an improvement in the degree of preferred orientation. R. Bacon and W.A. Schalman measured preferred orientation in terms of another parameter *Z*, which is orientation half width of the intensity of X-ray reflection along (002)

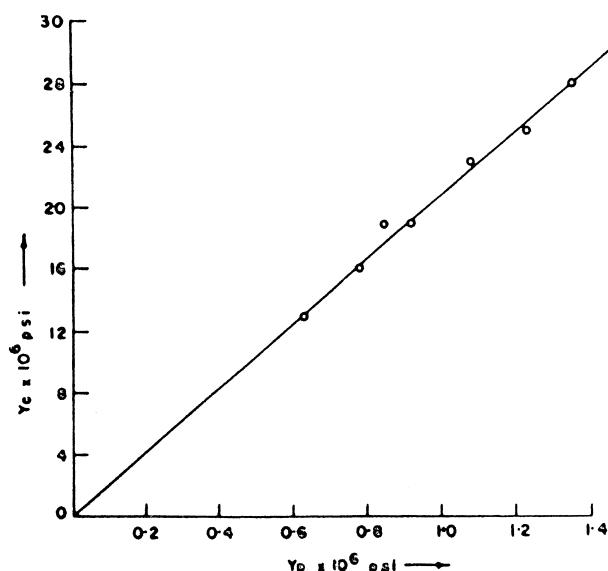


FIGURE 7 Relationship between the primary Young's modulus, y_p , and the Young's modulus of the resulting carbon fiber, y_c . [From Chari, S. S.; Bahl, O. P., and Mathur, R. B. (1981), *Fibre Sci. Technol.* **15**, 153.]

arc. This orientation half-width, Z , was found to vary between 25 and 5° for carbon fibers with Young's modulus between 35 and 560 GN/m².

The orientation of the crystallites in carbon fibers has been found to depend on the history of the precursor, the processing conditions, and the heat treatment temperature. In fact, a direct relationship has been observed between the primary Young's modulus of the precursor fiber, which in turn depends upon the preferred orientation of the crystallites, and that of the resulting carbon fiber (Fig. 7). The orientation of the crystallites is also enhanced by elongation of the fiber during one of the production stages because elongation renders the molecular chains closer to parallel to the fiber axis.

C. Porous Structure of Carbon Fibers

Carbon fibers have about 16–18% porosity, which exists in the form of needle-shaped pores with very small diameter. Small-angle X-ray scattering, which is generally used to estimate the size and the shape of the pores in carbon fibers, has shown that the predominant component in PAN- and rayon-based carbon fibers is a scattering of needle-like pores with diameters in the 1–2 nm range and length of about 20–30 nm. The pores have a preferred orientation of the needle axis parallel to the fiber axis and the angular distribution of their needle axis similar to that of graphitic layers.

The average diameter and the distribution of the pores is a function of the HTT, the processing conditions, and the nature of the precursor.

TABLE VI Pore Width and Internal Surface of PAN-Based Carbon Fibers from X-Ray Diffraction Studies

HTT (°C)	Pore width (nm)	Internal surface (Mm ⁻¹)
1000	0.86	1450
1400	1.16	1100
1800	1.28	980
1900	1.82	—
2000	2.54	485
2400	2.64	363
2800	2.80	340
2250 ^a	2.60	—
3000 ^b	4.0	280

^a Strain-graphitized.

^b Boron-doped.

The pore width increases while the internal surface decreases with increase in the HTT (Table VI). This is attributed to the elimination of some of the smaller pores with increasing graphitization. Hot stretching causes only a slight increase in the size of the pores, but the pores become more uniform. Doping with boron decreases the number of pores but increases their diameter. Porosimeter measurements have shown that the pores in the carbon fibers are inaccessible to helium but can be estimated by density measurements and by the adsorption of certain gases.

D. Structural Models

Several structural models have been proposed for carbon fibers. The first model based on X-ray diffraction and electron microscopic studies is the microfibrillar model (Fig. 8) proposed by W. Ruland and R. Perret. According to this model, the basic unit of the structure consists of ribbon-shaped monoatomic layers of sp²-type carbons with an average width of 6 nm and a length of several hundred nanometers. The ribbons display irregular contours which may contain holes. A certain number of these ribbons run parallel to form microfibriles, which have a preferred orientation parallel to the fiber axis. The ribbons show a high degree of parallel stacking and pass smoothly from one turbostratic domain to another, resulting in a turbostratic disorder. As the contours of the microfibrils are wrinkled, the space between two ribbons creates a number of needle-like voids, with length between 1 and 2 nm, and their orientation follows closely the orientation of the straight parts of the ribbons.

R.J. Diefendorf and E. Tokarsky, on the basis of TEM studies of HM carbon fibers proposed a wrinkled ribbon (basket weave) structure (Fig. 9), with ribbons varying in thickness from 13 to 30 layers and in width from 4 to 9 nm. The ribbons have almost zero amplitude and are essentially parallel to the fiber axis. In fact, the structure

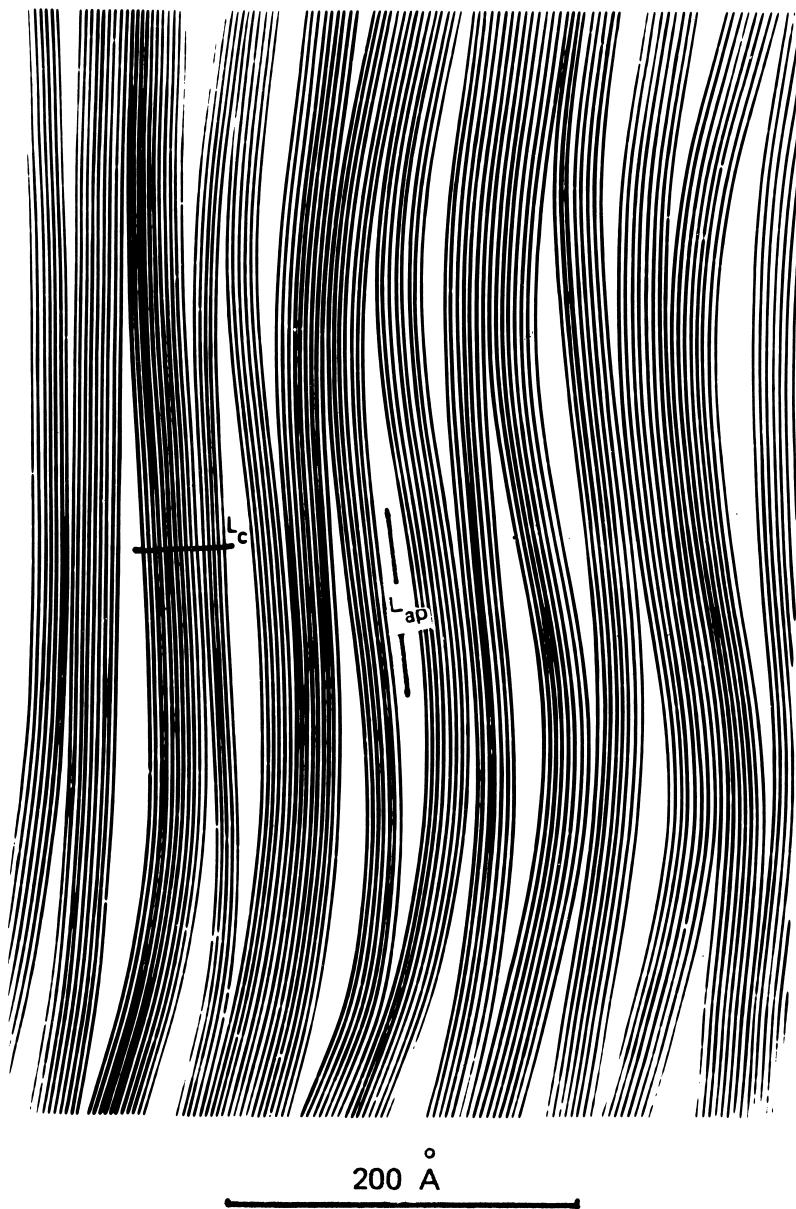


FIGURE 8 Schematic representation of fibrillar carbon fiber structure. [From Perret, R., and Ruland, W. J., *Appl. Crystallogr.* 3, 525.]

could be described as a wrinkled sheet. This sheet-like structure is shown by a TEM micrograph of HM carbon fiber. However, A. Oberlin et al., are of the view, based on their dark field examination, that a two-dimensional model for the longitudinal and transverse structure of HM and HT PAN-based carbon fibers could be better represented by Fig. 10.

E. Duplex Model

The cross-sectional structure of carbon fibers, which is important in constructing the three-dimensional model,

has been described as a circumferential-radial, or onion skin, or sheath-core model. R.N. Knibbs, while examining the end-on-end longitudinal section of different carbon fibers by optical microscopy identified three different types of fiber structures which differ in the orientation of the graphitic crystallites within the transverse plane (Fig. 11). Structures 11a and 11b are two-zone structures, while 11c is a single-zone structure. In 11a the central portion is almost isotropic, with circumferential orientation in the outer zone. Structure 11b exhibits radial preferred orientation in the middle zone and circumferential orientation in the outer zone, while structure 11c shows

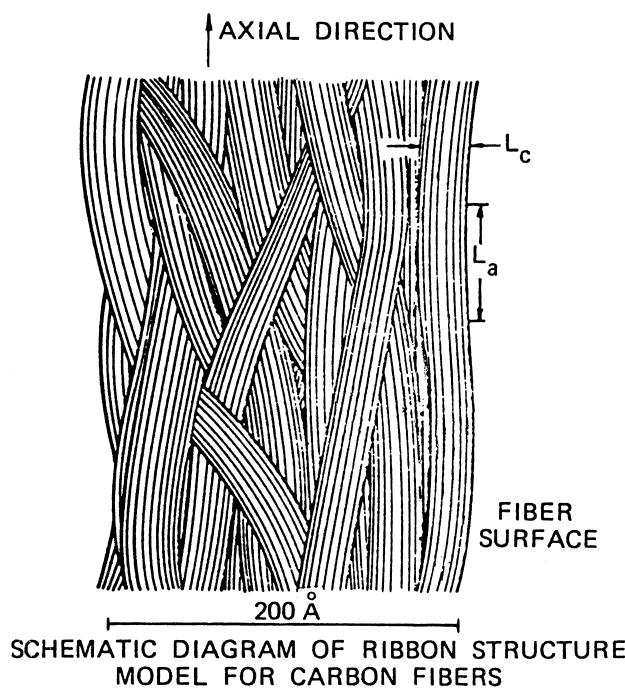


FIGURE 9 Basket weave model of carbon fiber structure. [From Diefendorf, R. J., and Tokarsky, E. (1975). *Polym. Eng. Sci.* **15**, 150.]

circumferential orientation throughout the cross section. The existence of these three structures was found to depend only on the difference in the precursor material or the processing conditions. SEM examination of oxygen plasma-etched carbon fiber surfaces showed that the central region was more etch resistant than the outer ring, supporting the duplex structure of carbon fibers. X-ray diffraction studies have shown that the outer sheath has a thickness of $1.5 \mu\text{m}$ and is separated by distinct annular region from the inner core, which is approximately $3 \mu\text{m}$ in diameter. In the case of carbon fibers graphitized at lower temperatures ($<2000^\circ\text{C}$), the transition between the sheath and core region was less distinct, and the transition region occupied a larger proportion of the fiber volume.

The cause of this sheath-core heterogeneous structure is the difference in the degree of stabilization in the outer and inner region of the carbon fiber. The fiber before carbonization and HTT has an almost completely stabilized outer region and only a partially stabilized inner region. These two regions form the sheath-core heterogeneous structure of carbon fibers.

F. Three-Dimensional Models

D. J. Johnson, on the basis of lattice fringe images of transverse sections of a number of different types of car-

bon fibers, has proposed the three-dimensional structure shown in Fig. 12. This model illustrates the difference in ordering and density of carbon ribbons between the peripheral zone and also the skin of the fiber which is in agreement with the behavior of the carbon fibers in many processes. A. Oberlin, however, has suggested an alternative model in which the layer planes have a variable transverse radius of curvature which decreases continuously from the surface to the center. All high modulus carbon fibers thus correspond to the model of transversally folded paper (crumpled sheet of paper).

VIII. SURFACE TREATMENT OF CARBON FIBERS

Carbon fibers with excellent mechanical properties can now be produced. When used without any surface treatment, however, they produce composites with poor inter-laminar shear strength (ILSS). This has been attributed

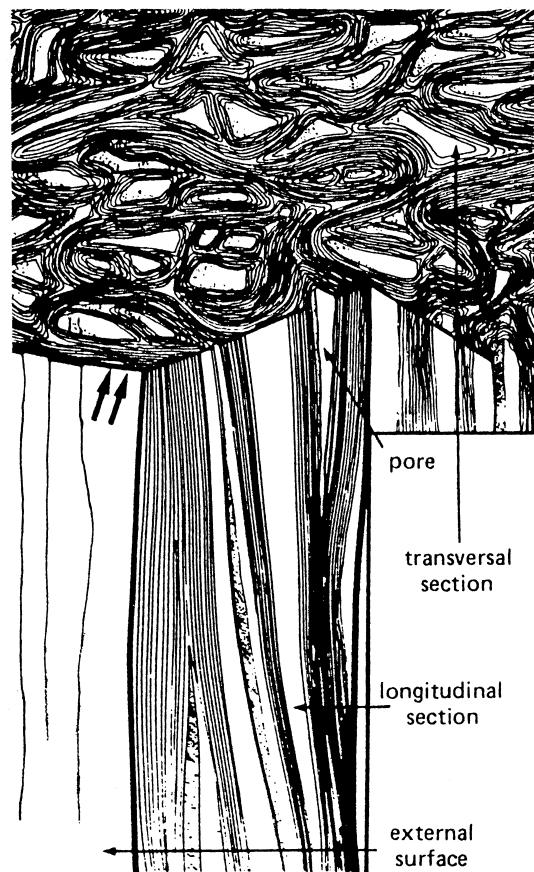


FIGURE 10 Two-dimensional representation of crystallite interlinking in carbon fibers. [From Oberlin, A., and Gurgaon, M., Conf. On Carbon Fibers, Japan.]

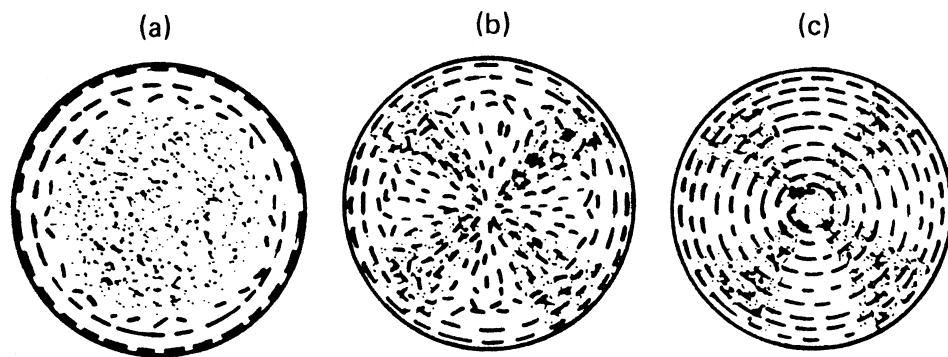


FIGURE 11 Schematic representation of carbon fiber structure: (a) isotropic center and oriented outer skin; (b) different orientation in the center and outer layer; and (c) one type of preferred orientation. [From Knibbs, R. H. (1971). *J. Microsci.* **94**, 273.]

to weak adhesion and poor bonding between the carbon fiber surface and the resin matrix. The interfacial bond between the fiber and the matrix can be enhanced either by enhancing the surface area, which provides more points of contact between the fiber and matrix, or by enhanc-

ing physicochemical interaction between the two. Thus, all carbon fibers are given a surface treatment which improves the fiber-matrix interfacial bond. The surface treatment should be such that it modifies the surface to make it more compatible with the matrix.

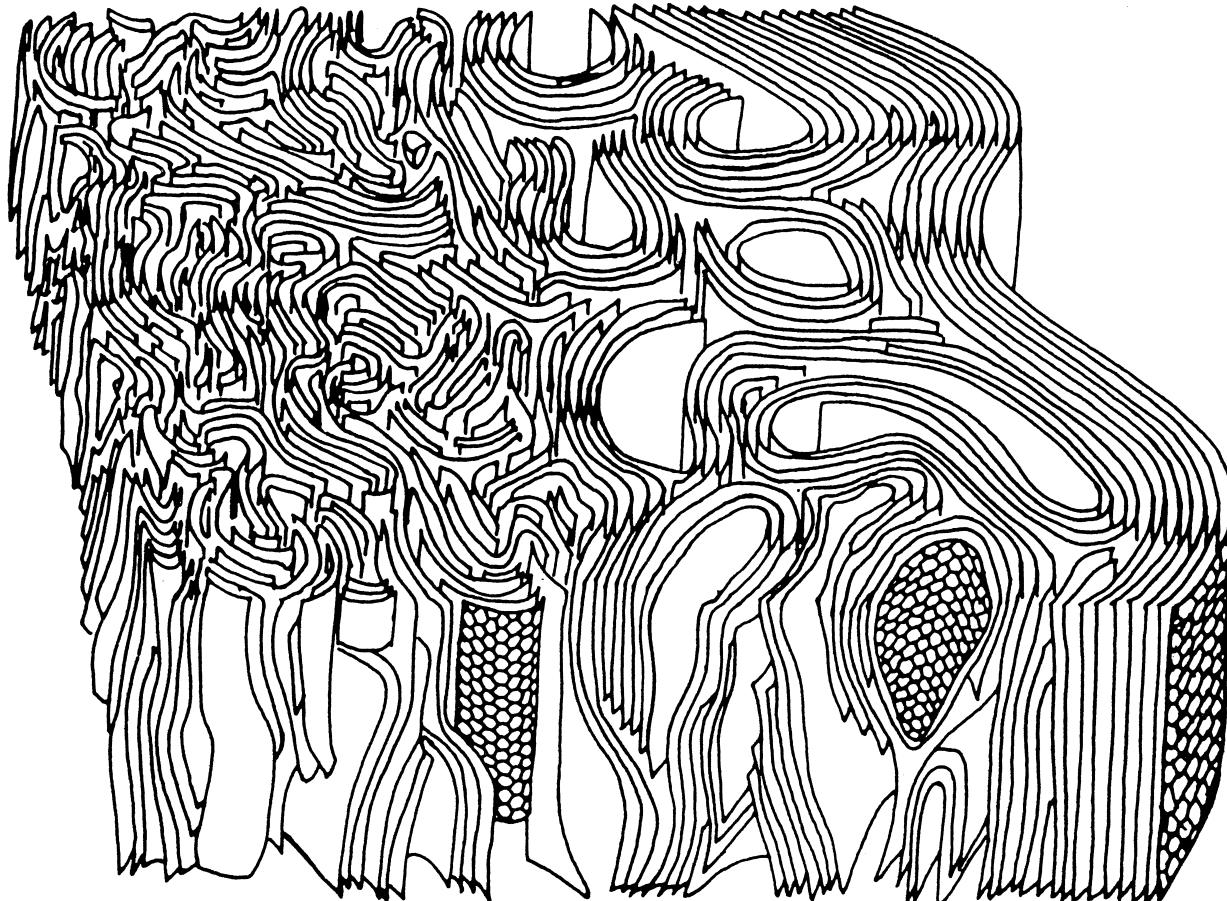


FIGURE 12 Three-dimensional model of carbon fiber structure. [From Bennet, S. C., and Johnson, D. J. (1979). *Carbon* **17**, 25.]

TABLE VII Surface Treatment of Carbon Fibers and Improvement of Composite ILSS

Treatment	Improvement in ILSS (%)
Gaseous oxidation (air, ozone, RF plasma)	10–15
Liquid-phase oxidation (HNO ₃ , NaClO, electrolytic)	100–200
Whiskerization (Si ₃ N ₄ , TiO ₄ , SiC)	200–300
Pyrolytic carbon coating (CH ₄ , FeC, SiC)	60–100
Polymer grafting	80–100

Several surface treatments have been examined. These may be classified into oxidative treatments and nonoxidative treatments. The oxidative treatment can be further divided into gas-phase oxidation at lower or at higher temperatures, liquid-phase oxidation carried out chemically or electrochemically, and catalytic oxidation. The nonoxidative treatments involve the deposition of a more active form of carbon such as the highly effective whiskerization, the deposition of pyrolytic carbon, or the grafting of polymers on the carbon fiber surface. Average improvement in composites in shear strength that have been obtained by the various treatment are recorded in Table VII.

Although a huge amount of work has been carried out by several groups of workers on oxidative surface treatment, the exact mechanism by which oxidation treatments improve the ILSS of composites is not completely elucidated and is still a matter of speculation. This is because virgin (untreated) carbon fibers are difficult to obtain. A considerable amount of evidence has accumulated, however, that the increase in surface area and the formation of surface acidic functional groups caused by the oxidative treatments improve bonding between fiber and the resin matrix. This tends to increase the wettability of the carbon fiber for the matrix and enhances the ILSS of the composites.

J.B. Donnet and coworkers have measured the polar and dispersive component of surface energy of a number of high-modulus and high-strength PAN-based and mesophase-pitch-based carbon fibers before and after several treatments by measuring contact angle using the two-liquid method and by using inverse chromatography. They are of the view that the improvement in adhesion between the fiber and the matrix is due not only to the creation of surface functionality but also to an increase in surface energy or surface roughness caused by the oxidation treatment.

Bahl and coworkers have reported a novel technique which helps in improving the ILSS of C/C composite by as much as 500%. In this technique commercial carbon

TABLE VIII Mechanical Properties of C/C Composites

Fiber used	HTT (°C)	Flexural strength (MPa)	F.M. (GPa)	ILSS (MPa)
T-800 (as such)	1000	100	48	4
	2600	130	50	5
T-800	1000	480	88	21
pyrocoated	2600	310	132	14

fibers are first given a thin coating (100 nm) of pyrocarbon by using a suitable hydrocarbon gas. C/C composites were prepared thereafter using coal tar pitch as the matrix precursor. The mechanical properties of C/C composites are given in Table VIII.

IX. PROPERTIES OF CARBON FIBERS

The physical properties of carbon fibers that make them versatile material for certain highly technical applications are their high specific strength and stiffness, light weight, low coefficient of expansion and low abrasion, good electrical conductivity, good vibration damping, biological compatibility, chemical inertness, and high fatigue resistance. The high modulus of all carbon fibers is due to the good orientation of the turbostratic graphite layer planes which constitute the material. These well-aligned layers of carbon atoms also give rise to good thermal and electrical conductivity. The stability of CFRP structure is assisted by a low coefficient of thermal expansion and further enhanced by excellent damping characteristics, chemical inertness, and biocompatibility.

A. Young's Modulus

Young's modulus is an intrinsic property and is governed by the orientation of the graphitic crystallites relative to the fiber axis. Thus, a sample of carbon fiber that had basal planes of crystallites lying within ±35° to the fiber axis shows a Young's modulus of 103 GN/m², while another sample with an orientation angle of ±10° shows a modulus of 410 GN/m². Therefore, any treatment, such as HTT or application of stress during one of the stages of processing, which could reduce this angle of crystallite orientation enhances the Young's modulus. Thus, it is now possible to achieve a Young's modulus of around 700 GPa in HM PAN-based carbon fiber and a modulus as high as 950 GPa in a mesophase-pitch-based carbon fiber. Unfortunately, high modulus fibers have a low strain-to-failure, whereas a high strain-to-failure is essential for many applications. Significant improvement have been made with high strength and intermediate modulus fiber types in

recent years, and strain-to-failure of around 2.5% can now be obtained with a commensurate tensile strength of around 7 GPa. HM pitch-based carbon fibers, which earlier showed only 0.3% strain-to-failure, are now available with more than 0.5% strain-to-failure.

B. Strength of Carbon Fibers

The strength of carbon fiber is a complex property and is measured under tension, bending, compression, and torsional shear. Strength depends on the type of precursor fiber, processing conditions such as HTT, preoxidation conditions, and the presence of flaws and defects. Tensile strength increases upon heat treatment, reaches a maximum value at about 1500°C, and decreases thereafter at higher temperatures. It is influenced by the nature of pretreatment atmosphere, the temperature, and the length of time of pretreatment. Heating PAN precursor fiber in air at 220°C up to 22 hr or more causes significant increase in strength, although heat treatment in nitrogen at the same temperature decreases the strength. Similarly, treatment in air at 215°C produces stronger fibers than treatment in oxygen.

The manufacturers' data for the strength and modulus properties of some carbon fibers are given in [Table IX](#). The most important carbon fiber for structural applications such as in aerospace and machinery support goods are the PAN-based intermediate heat-treated type (heat-treated at 1400–1500°C), although carbon fibers with much higher strength are available. This is due to high strain-to-failure for this fiber type. Thus the efforts of all carbon fiber manufacturers are directed to developing products with even higher strain-to-failure by means of further increase of strength but keeping the same intermediate modulus (250–300 GPa). The low strain to failure below 1% in PAN and below 0.5% in mesophase-pitch-based carbon fibers can be tolerated only in those applications where high Young's modulus is an absolute necessity, such as stiffness-controlled parts.

C. Electronic Properties

High modulus carbon fibers are composed of graphitic crystallites with their axes preferentially lying parallel to the fiber axis. The orientation is most suitable in enhancing the conductivity of carbon fibers. In addition, the geometric form of the filaments is more suitable for making satisfactory contacts and for engineering design of conducting materials. However, these materials have a certain degree of disorder which tends to decrease their electrical conductivity. This order can be reduced considerably by heat treatment of carbon fibers to a temperature as high as 3000°C, which creates larger domains of hexagonal structure. The electrical conductivity can also be considerably

enhanced by intercalation of carbon fibers. Mesophase-pitch-based or VGCF heat-treated to such temperatures as 3000°C have a highly graphitic structure and therefore show a high electrical conductivity, which can be further enhanced by intercalation. Benzene-derived carbon fibers heat-treated at 2800°C are highly ordered and, when intercalated with potassium, FeCl_3 , CuCl_2 , and AsF_5 exhibit a high conductivity compared to highly oriented pyrolytic graphite (HOPG) ([Table X](#)). The most metallic values are obtained with AsF_5 intercalations, the value being almost similar to that of copper. Thus, the VGCF, which on intercalation produce highly conducting organic materials, have potential as conductors in the aerospace industry. A great effort is still needed to make these intercalated materials withstand the influences of heat, air, and water.

D. Thermal Properties

With the steep rise in rate of industrialization there is an urgent need to develop materials for thermal management. Theoretically speaking graphite is an ideal material for this purpose. [Table XI](#) gives data for graphite.

These are highly lucrative properties and it is possible to realize them in actual components/products only through use of carbon fibers. [Table XII](#) below gives thermal conductivity data, along with mechanical values and crystallite size of some of the grades of carbon fibers. Polyacrylnitrile-based carbon fibers (though with very high modulus) show minimum value of thermal conductivity, whereas, pitch-based carbon fibers give thermal conductivity varying from 180 to 900 W/mK. There is a strong dependence of thermal conductivity on crystallite size La. Polyacrylnitrile gives hard carbon whereas pitch gives soft carbon, and this explains higher values of La in case of pitch and hence higher values of thermal conductivity.

X. CARBON FIBER DEMAND-PATTERN

There are two angles of looking at this: one is the total global demand and second is the changes in the demand pattern. It is in fact the demand-pattern changes that can help us predict the future of the materials. [Fig. 13](#) gives actual demand up to 1998 and is extrapolated up to 2005. In the early 1980s, demand ratio (industrial:sports:aircraft:space) was 18%:37%:45%. As it should be, total requirement is being driven by specialized applications, i.e., in aircraft and aerospace sectors. The pattern changed to 25:40:35 by 1990, which is not substantially different from the pattern in 1985. In 1998, the demand-pattern is lead (48%) by the industrial sector which is indeed a very healthy sign for general acceptance of any material. If the demand-pattern continues, it is predicted that by the year 2005, total demand for industrial

TABLE IX Mechanical Properties of Carbon Fibers (Manufacturers' Data)

Manufacturer	Fiber	Young's modulus (GPa)	Tensile strength (GPa)	Strain-to-failure (%)
PAN-based high modulus (low strain-to-failure)				
Celanese	Celion Gy-70	517	1.86	0.4
Hercules	HM-Smagnamite	345	2.21	0.6
Hysol Grafil	Grafil HM	370	2.75	0.7
Toray	M 50	500	2.50	0.5
-do-	M 55J	540	3.63	0.7
-do-	M60 J	600	3.7	0.6
-do-	M70 J	700	3.5	0.5
PAN-based intermediate modulus (intermediate strain-to-failure)				
Celanese	Celion 1000 IM-6	234	3.24	1.4
Hercules		276	4.40	1.4
Hysol Grafil	Apollo IM 43-600	300	4.00	1.3
Toho Beslon	Sta-grade Besfight	240	3.73	1.6
Union Carbide	Thornel 300 M 30	230	3.10	1.3
Toray		294	3.92	1.3
PAN-based HT (high strain-to-failure)				
Celanese	Celion ST AS-6	235	4.34	1.8
Hercules		241	4.14	1.7
Hysol Grafil	Apollo HS 38-750	260	5.00	1.9
Toray	T-800	300	5.70	1.9
-do-	T-1000	294	7.06	2.4
Mesophase-pitch-based				
Union Carbide	Thornel P-25	140	1.40	1.0
	P-55	380	2.10	0.5
	P-75	500	2.00	0.4
	P-100	690	2.20	0.3
	P-120	820	2.20	0.2
	K-1100	940	3.1	0.33
Osaka Gas	Donacarbo F-140	140	1.80	1.3
	F-600	600	3.00	0.5
Tonen Corporation	FT 700	700	3.3	0.45
	FT 500	500	3.0	0.60
Nippon Oil Company	XN 50	490	3.2	0.65
	XN 70	690	3.3	0.48
Mitsubishi Chemical Corporation	Dialed KB 710	636	2.9	0.46
	K 63712	636	2.66	0.42
	K 13 A 10	790	2.8	0.35
	K 63 B 12	890	3.03	0.34

and sports sectors put together would be around 90% which is a positive indicator that carbon fibers, as a new material, is approaching maturity.

XI. FURTHER AGENDA

Figure 14 gives an overall view about how the property improvement in PAN- as well as pitch-based carbon fibers has taken place over the years. In case of HT carbon fibers, tensile strength of PAN-based fibers has improved

from 2 GPa in 1968 to about 7 GPa in 1998. On the other hand, Young's modulus registered an increase from 200 GPa to about 700 GPa. We are moving upward on the X-axis or toward the Y-axis. Ideally speaking, however, the arrow in the diagram indicates the desired direction of improvement implying simultaneous improvement in both TS as well as Young's modulus. Improvement of properties over the years shows there are two areas, namely precursor quality (particularly in terms of defects, etc.) and processing including preoxidation, which control the final mechanical properties of carbon fibers. Defects in

TABLE X Room Temperature Electrical Resistivity for Intercalated Benzene-Derived Carbon Fibers

Fiber	Resistivity ($\mu\Omega\text{cm}$)
Original fiber	80
Fiber intercalated with:	
Potassium	6–9
FeCl_3	5.1 (HOPG: 4.3–9.5)
CuCl_3	9.4
AsF_5	2–4 (HOPG: 1.6–2.5)

the polymer PAN precursor fibers as well as uniformity in preoxidation and carbonization could be controlled if one can imagine a smaller diameter precursor. The average diameter of first generation (TS 2 GPa) carbon fibers is around 7–8 μm and that of present generation (TS 7 GPa) is around 4.5–5 μm . This indicates that manufacturing companies have tried and succeeded in spinning thinner fibers thus improving preoxidation uniformity and hence the mechanical properties of final carbon fibers. The situation is altogether different as far as pitch-based carbon fibers are concerned. Carbon fibers (first generation) with YM ranging from 400 GPa to around 900 GPa and TS of around 2 GPa only were available. Pitch yields soft

TABLE XI Thermal Characteristics of Single Crystal Graphite

	Along a axis	Along c axis
Thermal conductivity W/mK	2000	10

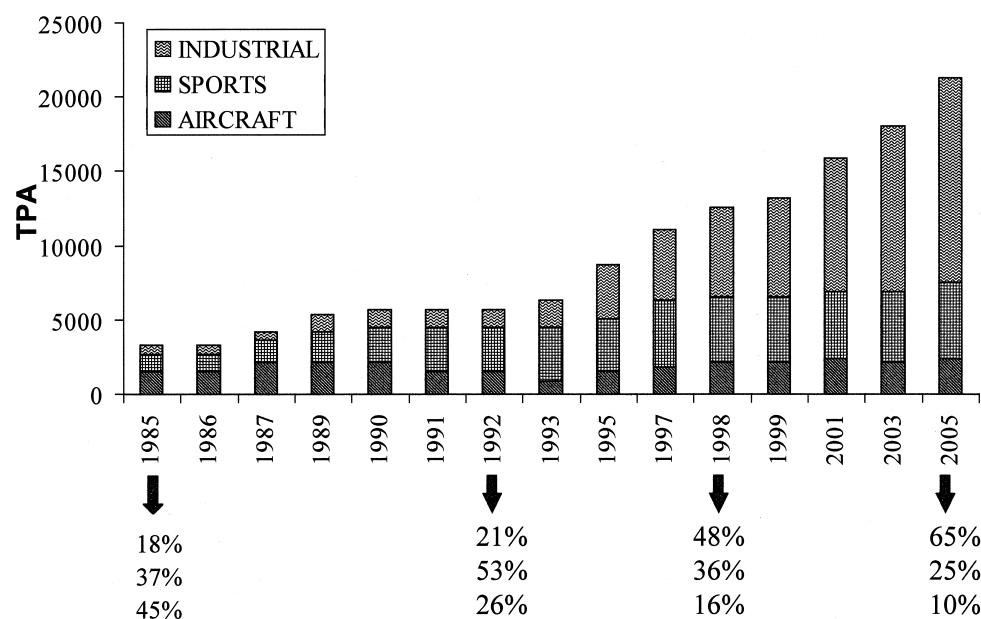


FIGURE 13 Changes in carbon fiber demand-pattern.

TABLE XII Thermal Conductivity Values for PAN- and Pitch-Based Carbon Fibers

Carbon fiber type	Tensile strength (GPa)	Young's modulus (GPa)	Thermal conductivity (W/mk)	Crystallite size La (\AA)
M 60 J (PAN) Torayca	3.8	600	100	10
P 75 (pitch) Amoco	2.0	525	180	—
P 100 (pitch) Amoco	2.2	760	510	60
K 1100 (pitch) Amoco	3.1	940	900	100
K 13 D (pitch) Mitsubishi	3.7	940	800	55
Copper	—	—	400	—

carbon which attains graphitic structure easily thus offering a possibility of failing in shear. This is the reason why TS of pitch-based carbon fibers could not be improved beyond 2 GPa, whereas YM could be improved at will almost reaching theoretical numbers. In the recent past, significant improvements in pitch making have taken place. Companies have succeeded in introducing substantial cross-links hence reducing chances of failure of carbon fibers in shear. Tensile strength of present generation pitch-based carbon fibers is in the range of 3.5–4.0 GPa. With these improvements, pitch-based carbon fibers are competing with HM. PAN-based carbon fibers along with the whole scenario are liable to change in the future.

XII. APPLICATIONS OF CARBON FIBERS

The driving force for developing carbon fibers has been their use in the aerospace sector primarily because of their

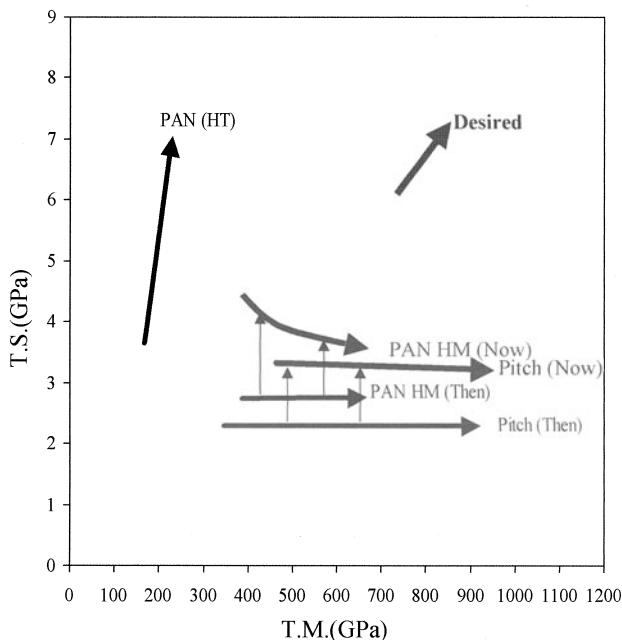


FIGURE 14 Changes in mechanical properties of PAN- and pitch-based carbon fibers.

high stiffness and low weight. Besides these mechanical advantages there are other promising properties which are offered for new applications of carbon fibers in composites, such as low thermal expansion (near zero) surprising high fatigue resistance (better than any metal), extremely good corrosion resistance, etc.

Carbon fibers are used in composites with a lightweight matrix, generally an epoxy resin and occasionally a polyester or a polyamide. For high-temperature applica-

tions, they are also used in composites with certain pitch and other refractory materials, which on pyrolysis produce carbon matrices. The three broad sectors of carbon fiber application are the high-technology sector such as aerospace and nuclear engineering, industrial, and sports. **Table XIII** shows the global market of carbon fibers from different manufacturers for the last three years. The requirements for these sectors, however, are entirely different. The large-scale use of carbon fibers in aerospace and aircraft is driven by high performance and fuel efficiency, while use in general engineering and surface transportation is determined by cost constraints, high production rate requirements, and generally less critical performance needs. This leads to different considerations regarding material forms, acceptable matrices, potential manufacturing methods, and on-line quality control requirements.

A. Aerospace Applications

Aerospace has been a primary influence on the development of carbon reinforced composites and constitutes the most significant market sector. The oldest flight-tested components which have performed well under the rigors of in-service use are rudder twin tabs on the jet Provost, airbrakes on the Vulcan, a flat operating torque tube on the Andover, a door strut on the Trident, and ferry wing tip on the Harrier. All the components showed absence of corrosion and foreign impact object damage and apparent resistance to aircraft fluid in service. **Figure 15** illustrates some of the areas in which the advanced carbon fiber composites are considered competitive with metals and where design specifications on deflection limits, buckling, and dynamic response cannot be met without them. In addition, the floor

TABLE XIII Global Market Status of Carbon Fibers (Capacity Quoted ton/year)

Producer	Country	Capacity 1997	Capacity 1998	Capacity 1999
Aerospace applications (1000–24000 monofilament)				
BP—Amoco	U.S.	2100	2100	2100
Hexcel (ex. Hercules)	U.S.	2000	2400	2400
Mitsubishi Rayon (without Grafil)	J. (U.S.)	1700	3200	3200
Taiwan Plastics (Formosa Plastics)	Taiwan	360	1000	1000
Toho (without Tenax)	J. (D)	3700	5100	5100
Toray (without Soficar)	J. U.S. (F)	3700	5600	7400
Total		13560	19400	21200
Industrial applications (40,000–320,000 monofilament)				
Acordis (AKZO Nobel)	U.S.	2100	2300	3600
Aldia	U.S.	0	600	1200
SGL Carbon Group (without RK Carbon)	D (GB)	850	1850	2850
Zoltek	U.S., Hun.	3800	8900	16100
Total		6750	13650	23750

[From Korinek, Z. (1999). Acta Montana IRSM AS CR, Series B, **9**, 19–23.]

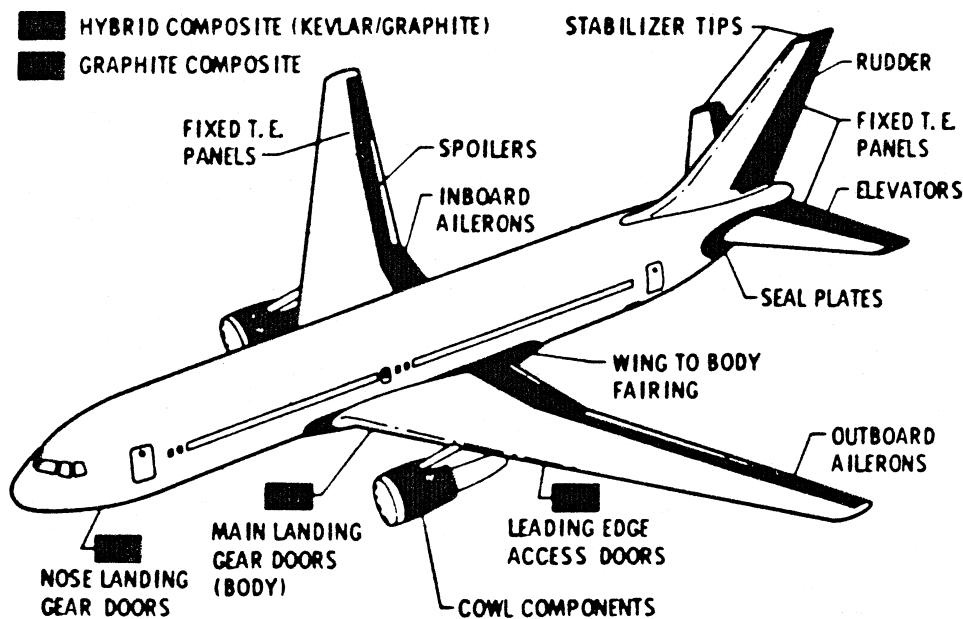


FIGURE 15 Air frame applications of carbon fiber composites. [From Thomas, D. K. (1983). *Plast. Rubber Int.* **8**, 53.]

decking of the Boeing 747 and the weapons bay door of the B-1 bomber are made of carbon fiber reinforced plastics (CFRP). Carbon fiber is being used now in primary structures of both Boeing 777 as well as Airbus A 320, etc.

In the case of helicopters, the most significant areas of application of CFRP are the main and tail rotor blades, where the prospects of improved fatigue resistance over all-metal blades have been recognized and rapidly exploited.

The largest piece of CFRP structure in service at the moment is in the payload bay doors of the NASA space shuttle, which are 60 ft long and 1500 square ft in area. The use of CFRP has resulted in an estimated weight saving of 23% over the same structure in aluminum alloy. The most widespread use of CFRP is in the advanced Harrier AV-8B jump jet, where advanced composites are expected to account for about 26% of the frame structure. Weight saving is also important in airborne equipment, such as a Plassey telescope cabin, which is a complete radar installation carried to forward areas on a helicopter that shows a weight saving of 26% when fabricated from CFRP.

There is exploding worldwide demand for mobile communications and high-speed broad band multimedia services. In turn, this will create an unprecedented need to place thousands of satellites and vehicles used to launch them, thus creating a wide use of carbon composites. It is projected that during the first decade of the new millennium a total of 1200–2000 satellites will be launched. Further, most transportation systems place a premium of some type on weight. In the case of satellites and launch

vehicles, a premium of \$15000–25000/kg as compared to \$200–1000/kg in the case of aircraft, is well recognized. As a matter of fact, satellites are in a cost class of their own as far as need to save weight is concerned, since the payload constitutes only 1–3% of the total weight of any launch vehicle.

The exceptional thermophysical properties coupled with their strength and stiffness are finding wide applications in the form of carbon-carbon composites (carbon fiber reinforced carbon; (CFRC). These applications include reentry vehicle heat shields, temperature ducting systems, nuclear rocket engines, rocket nozzles, heat exchange tubes of helium-cooled nuclear reactors, and aircraft brakes. When made from carbon-carbon composites, aircraft brakes are found to have remarkably stable frictional characteristics to temperatures well above 2000°C. Carbon is stronger at 2000°C than at room temperature, while steel melts at 1600°C. The carbon fibers also provide the resistance to thermal shock that is so important in this application. Goodyear has CFRC air brakes weighing 30 kg (compared with 100 kg for steel brakes) and found them to be three to five times better than their steel counterparts in actual number of landings before burning out. Thus, carbon-carbon composite brakes provide more than weight saving alone as they also perform better.

The high-temperature properties of CFRC have also been used in making molds for metallurgical hot forging processes in the super plastic stage and for hot molding in powder metallurgy. The low coefficient of linear expansion of carbon fibers is used to design structures with very

low linear or planar thermal expansion, such as telescopes and aerospace antennae.

B. Applications in Sports

Superior specific strength and stiffness with good fatigue resistance and damping make CFRP a versatile material for certain sports and leisure-time products such as golf club shafts, fishing rods, ski poles, racing bicycle frames, vaulting poles, racing cars, and bows and arrows. Reinforcement with carbon fibers results in a lighter shaft, permitting proportionately more weight than usual in the head of the golf club. Total consumption of carbon fibers in the annual manufacture of golf shafts alone is around 3000 MTPA which is very significant.

C. Application in the Chemical Industry

Because they are chemically inert to a range of chemical reagents, carbon fibers in the form of mat, felt, or cloth impregnated by a resin matrix can be used as a protective lining for chemical plants and anticorrosion equipment, including bulk liquid chemical transport. Carbon fiber reinforced plastics can also be used for making blades for rotary compressors and in the construction of gas centrifuges for uranium enrichment. FRP and metals react with fluorine in uranium hexafluoride in the uranium enrichment and therefore require a protective lining. No such coating is needed in the case of CFRP, in addition to resulting in weight saving.

D. Electrical Applications

Good electrical conductivity of carbon fibers makes them attractive fillers in polymers for making conducting composites which can be used for electromagnetic interference shielding, aircraft surfaces tolerant to lightning strikes, electrostatic dissipation for the protection of microcircuitry in the computer field, and for large conducting space structures. The conductivity of the composite can be controlled by proper selection of carbon fiber and the matrix material to meet different applications.

E. Biomedical Applications

Carbon is well known to have the best biocompatibility. It is compatible with blood, soft tissue, and bones. This biocompatibility coupled with the specific strength and pliability of carbon fibers makes them useful surgical implants for regenerating tissues and as replacements for damaged or absent tendons. Carbon fibers have exciting applications as components of bone plates, hip joint prostheses (Fig. 16), ligaments, and hydraulic motors for artificial

heart implants. Carbon fiber reinforced plastics are also being used successfully to produce the main load-bearing structure of a new artificial limb for the lower extremity. The high permeability of carbon fibers to X-rays is made use of in the construction of the patient's couch in computerized tomography equipment and for cantilever tables in therapeutic X-ray equipment.

G.A Illizarov of erstwhile USSR had, during World War II, invented a device known as Illizarove fixator wherein stainless steel rings were used to fix or join broken bones etc. Stainless steel has been replaced with carbon fiber composite rings which score heavily over stainless steel rings in respect of the following:

1. These are elastically rigid over the whole range of loads and avoid the plastic deformation exhibited by stainless steel rings.
2. These are transparent to X-rays and thus allow better monitoring of the healing process.
3. Because of their linear load deformation response, these can be reused.

These carbon rings have been successfully adapted for curing polio patients (Fig. 17), and a new era in the management of handicapped (polio) patients which hitherto was not possible has thus been created.

F. General Engineering Applications

Several general engineering components such as bearings, gears, cams, fan blades, bolts, screws, nuts, and washers can be made from carbon fiber reinforced composites. They can also be used in the field of apparatus construction in the form of tubes, sleeves, and other auxiliary aids. In the case of textile machinery, CFRP is being used extensively to make such components as picking sticks, reciprocating guide bars, push rods, and slays in weaving looms. The reciprocating guide bar which reciprocates longitudinally at the rate of 750 cycles per minute when made from CFRP has increased the speed by about 10% and has doubled the life of the bar.

G. Adsorption

Traditionally used activated carbons were either in the form of powder or granules. With the availability of carbon in the form of carbon fibers, a third form, known as activated carbon fibers, (ACT) has successfully been developed.

Activated carbon fibers are an adsorbing material which outperforms granulated active carbon (GAC) because

- it possesses a highly developed micropore structure, which characterizes huge adsorption capability and

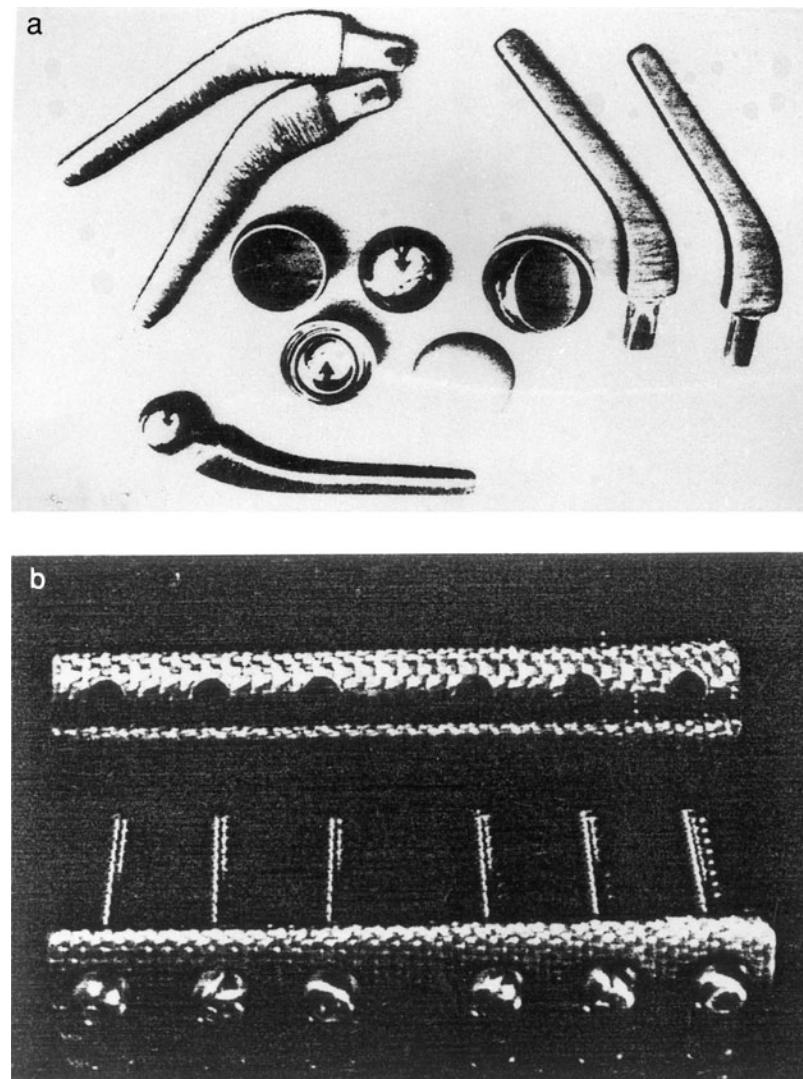


FIGURE 16 Carbon fiber composite for prosthetics and bone joints. [From Sepcarb Bulletin 1981.]

capacity that is maintained well under the condition of low concentration;

- both adsorption and desorption are very fast, especially complete desorption can be achieved under simple conditions;

- it is extremely easy to fabricate into various forms of strand, nonwoven fabric, woven fabric, paper, honeycomb, and other type of configurations to meet various needs;

- through the process of impregnation, ACF can be equipped with added catalytic, chemisorbic, and bactericidal properties;

- the adsorption speed of ACF can be from 10 times to several hundred times than that of GAC.

Major applications of ACF are in solvent recovery systems including cleaning the vapors and separating the vapors as well as solvent recovery and air conditioning.

The odor compounds emitted from the human body and the remains of plants and animals include ammonia, acetone, isovaleric, and diethylamine and can be removed by using ACF. In addition, ACF also removes particles and moisture existing in the air, particularly aromatics which may cause cancer. Activated carbon fibers can be applied to the air cleaning apparatus in hotels, hospitals, offices, trains, and airplane cabins as well as at homes.

- *Waste water treatment:* ACF is being effectively used for waste water treatment in pharmacy, printing, oil refining, paint industry, etc.
- *Water purification:* ACF is an ideal material as purifying medium in the production of drinking water and super pure water for industrial use.

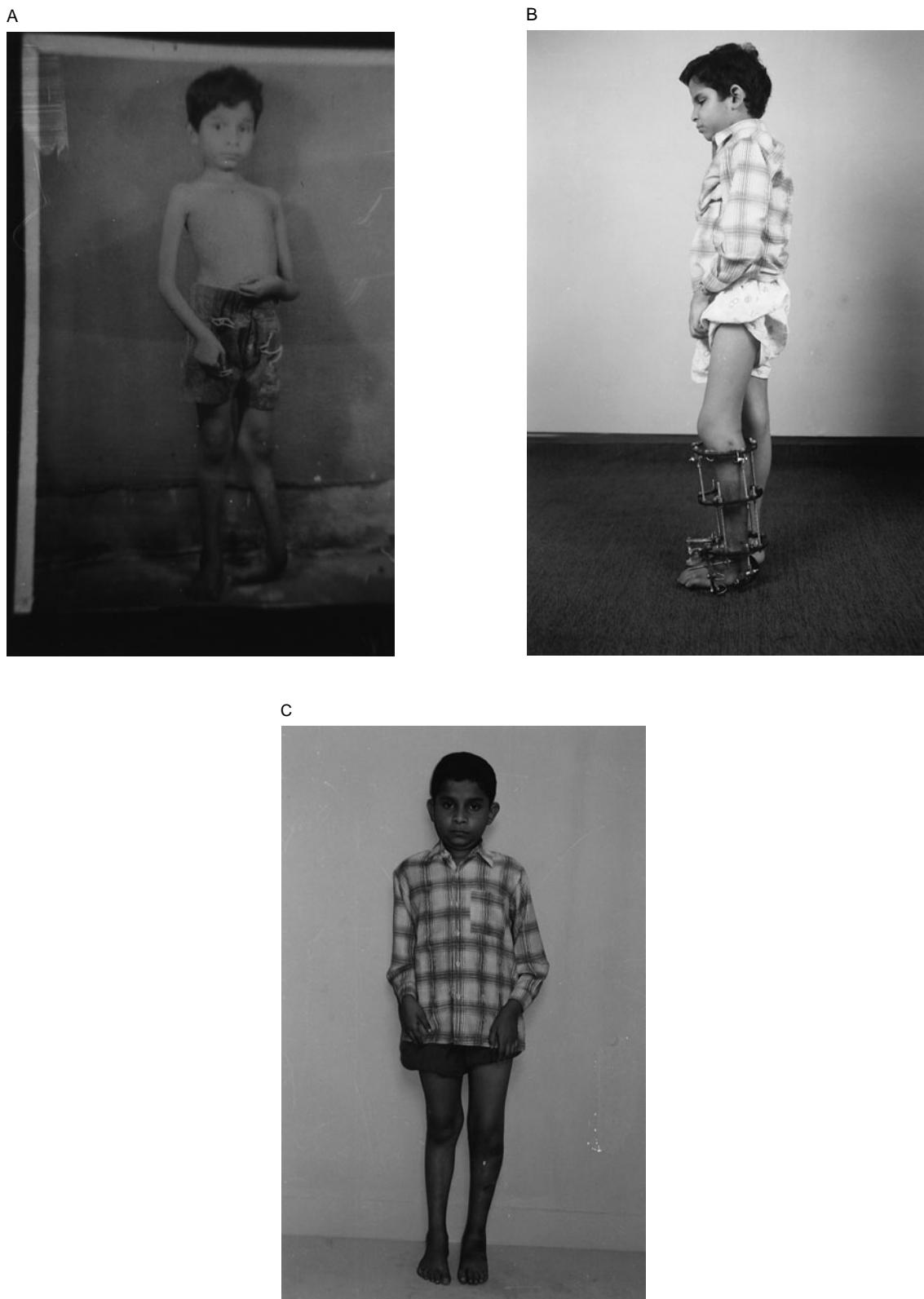


FIGURE 17 Polio affected child. (a) Before any treatment. (b) Fitted with carbon-fiber-based Ilizarov fixator. (c) Completely cured of polio. [Carbon composite rings designed and made by Bahl, O. P., Mathur, R. B., Dhami, T. L., and Inderjeet Singh and the patient operated upon by Agarwal, R. A. at Agarwal Orthopaedic Hospital, Gorakhpur, India.]

- *Chemical protection:* ACF possess outstanding capability and reliability to toxic agent including war chemical agent of organic phosphorous. It can also efficiently shield out phosgene.
- *Medical and biological:* ACF is commonly used as the adsorption medium in artificial kidneys as well as for making bandages and various kinds of bacterial-killing products.

It is firmly believed that carbon fibers in ACF form will remain an indispensable material at least in the coming five decades in all walks of life for environmental protection, industrial protection, and household applications.

Finally, the growth of the carbon fiber market was sharply slower than expected in recent years and even if the recent increase in uses are very positive, many carbon fiber suppliers have left the market in the last 15 years. Values for the market should be taken with care with the future remaining still quite difficult to assess.

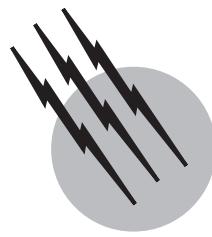
SEE ALSO THE FOLLOWING ARTICLES

BIOPOLYMERS • COMPOSITE MATERIALS • METAL MATRIX COMPOSITES • POLYMERS, MECHANICAL BEHAVIOR

- POLYMERS, THERMALLY STABLE • SANDWICH COMPOSITES

BIBLIOGRAPHY

- Bahl, O. P., Mathur, R. B., and Dhami, T. L. (1984). *Polym. Eng. Sci.* **24**, 455.
- Chari, S. S., Bahl, O. P., and Mathur, R. B. (1981). *Fibre Sci. Tech.* **15**, 153.
- Delmonte, J. (1981). "Technology of Carbon and Graphite Fiber Composites," Van Nostrand Reinhold, Princeton, New Jersey.
- Donnet, J. B., Wang, T. K., Rebouillat, S., and Peng, J. C. M. ed. (1998). "Carbon Fibers," Marcel Dekker, New York.
- Donnet, J. B., and Bansal R. C. (1988). "Carbon Fibres," Marcel Dekker, New York.
- Donnet, J. B., Dhami, T. L., Dong, S., and Brendle M. J. (1987). *J. Phys. D: Appl. Phys.* **20**, 269.
- Ehrburger, P., and Donnet, J. B. (1985). "High Technology Fibers," (M. Lewin and J. Preston, eds.), Marcel Dekker, New York.
- Fitzer, E. (1985). "Carbon Fibers and Their Composites," Springer-Verlag, New York.
- Fitzer, E., Frohs, W., and Heine, M. (1986). *Carbon*, **24**, 387.
- Guigou, M., Oberlin, A., and Desarmot, G. (1984). *Fibre Sci. Tech.* no **1** **20**, 55.
- Mittal, J., Mathur, R. B., and Bahl, O. P. (1997). *Carbon* **35**, 1713.
- Mochida, I., Shimizu, K., Otsuka, H., Sasaki, Y., and Fujiyama, S. (1990). *Carbon* **28**, 311.



Ceramics

Shigeyuki Sōmiya

Tokyo Institute of Technology

- I. Electronic Ceramics
- II. Structural Ceramics
- III. Ceramics for Biomedical Applications

GLOSSARY

Composite electrolyte Mixture of ionic conductors and insulating materials.

Concentration cell Cell whose electromotive force is generated by the difference in concentration of electrochemically active components.

Connectivity Key feature in designing the microstructure of PZT-polymer composites. Each phase in a composite may be self-connected in zero, one, two, or three dimensions. Using an orthogonal axis system, for diphasic composites, there are 10 connectivities designated as 0–0, 1–0, 2–0, 3–0, 1–1, 2–1, 3–1, 2–2, 3–2, and 3–3.

Coulomb efficiency Percentage of electricity used only to oxidize or reduce (deposit or dissolve) ionic species electrochemically.

Curie temperature (T_c) Temperature marking the transition between ferroelectric and nonferroelectric properties of a material.

Curie–Weis law Curie–Weis law is expressed as $\varepsilon/\varepsilon_0 = C/(T - \theta)$, where ε is the small-single absolute permittivity, ε_0 the permittivity of free space, C the Curie constant, and θ the Curie temperature.

Decomposition voltage Voltage at which an electrolyte starts to oxidize or reduce electrochemically.

Dielectric constant Property that determines the quantity of electrostatic energy stored per unit volume for unit potential gradient.

Dielectric loss Time rate at which electric energy is transformed into heat in a dielectric when subjected to a changing electric field.

Electrode Terminal or surface at which electricity passes from an electron-conductive medium to an ion-conductive medium or vice versa.

Electrode potential Voltage existing between an electrode and the electrolyte with which it is in contact.

Electrolyte Liquid, paste, solid, or other conducting medium in which the flow of electric current takes place by migration or diffusion of ions.

Electromotive force Difference in electrode potentials produced in an electrolytic cell.

Energy trapping In a piezoelectric plate with partial electrodes on each surface, vibrational energy is trapped in the electrode region, when the piezoelectric materials satisfy the particular conditions corresponding to the vibration mode.

Exhaust port liner Duct that receives exhaust gas from exhaust valve.

Gas turbine Engine having rotary blades driven by a flowing hot gas.

Glow plug Plug in a diesel engine cylinder that ignites injected fuel on its hot surface.

Insulation resistance Resistance offered by insulation to the flow of current from an impressed direct voltage.

Interdigital transducer (IDT) Double comb-shaped structure of interdigitate metal electrode teeth applied to the surface of a substrate that transforms electrical energy into acoustic energy or vice versa by means of the piezoelectric effect.

Intermediate frequency (IF) Frequency after a frequency conversion before demodulation in a superheterodyne receiver.

Ion Charged atom or group of atoms. A negative ion (anion) has gained one or more extra electrons, whereas a positive ion (cation) has lost one or more electrons.

Ionic conductivity Measure of the ability of an ionic conductor to conduct electric current by ion transport, expressed as S/m or S/cm.

Ionic transference number Percentage of electric current transported by ions through an electric conducting medium.

Mechanical quality factor (Q_m) Factor representing the efficiency of mechanical vibration, expressed by the equation: $f = 2\pi^*A/B$, where A is time average of total energy stored in a material, and B is the energy dissipated per cycle at a given frequency due to the internal friction of the material.

Ohmic contact Contact between two materials across which the potential difference is proportional to the current passing through.

1-3 Composite Composite of polymer and piezoelectric ceramic fibers or rods, which is aligned to the poled direction and embedded in a polymer matrix.

Piezoelectric ceramics Ceramics that are polarized in the direction of high direct current field applied to them. They distort in an electric field, vibrate when an alternating current signal is applied, and generate electric charges under mechanical stress.

Polyvinylidene fluoride (PVF₂) Typical piezoelectric polymer material.

Potential barrier Barrier in which the charge carrier charge density is insufficient to neutralize the net fixed charge density of donors and acceptors.

Precombustion chamber Small tubelike component having a cavity and opening, used to evaporate injected fuel to mix gas with air in diesel engines.

PZT Lead zirconate titanate [Pb(Zr,Ti)O₃], typical piezoelectric ceramic materials.

PZT-5H Typical PZT for pick-up elements and audio tone transducers.

PZT-4 Typical PZT for ultrasonic transducers.

Rare earth Elements numbered 57 through 71 in the periodic table of elements.

Reciprocating engine Engine having piston and cylinder.

Resistivity Material property equal to the value of an electric field at a point in the material divided by the conduction-current density at the same point.

Rocker arm Component to open and close intake and exhaust valves with a mating cam.

Spontaneous polarization (P_s) Magnitude of the polarization within a single ferroelectric domain in the absence of an external electric field.

Temperature coefficient of resistance (A) Thermistor characteristic expressing the change in electric resistance (of positive temperature coefficient in this section) to a change in temperature.

Thermistor constant (B_1 and B_2) Thermistor characteristic expressing the change in electric resistance of negative temperature coefficient due to a change in temperature.

3-1 Composite PZT-block with holes drilled through to the poled direction of the PZT.

3-3 Composite Three-dimensionally interconnected PZT ceramics with connected pores filled with polymer.

Traveling wave Wave in which energy is transported from one part of a medium to another.

Turbocharger Turbine to charge compressed air to engine, driven by exhaust gas.

Ultrasonic Pertaining to signals involving frequencies just above the range of human hearing, hence above about 20,000 Hz.

0-3 Composite Piezoelectric composite consisting of a polymer matrix loaded with ceramic powder.

MAJOR ADVANCES continue to be made in ceramic science and technology. Progress has been made in both traditional and advanced technical ceramics. This article focuses on high-technology ceramics, divided into two fields: electronic and structural.

The field of electronic ceramics alone is too broad to be covered by a single author. In this article the field has been broken down into seven areas that have undergone major advancement. Omitted because of space is a discussion of magnetic ceramics; and because of the rapid changes in superconducting ceramics, any statement would be out of date before publication.

In the field of structural ceramics the most remarkable developments have been made in automobile parts. Ceramic elements are now replacing metallic ones for reasons of temperature- and stress-endurance among others.

This article presents the 1988 state of the art science and technology in ceramics. It is hoped that it will go far in promoting the understanding and development of advanced technical ceramics.

I. ELECTRONIC CERAMICS

A. Ionic Conductors

Masatomo Yonezawa*

Ionic conductors are defined as high-conductivity solid electrolytes having a high ionic conductivity ($10^{-1} \sim 10^{-5}$ S/cm) at fairly low temperatures below their melting points. They are also called fast ionic conductors or super ionic conductors when the conductivity is higher than 10^{-3} S/cm. Ionic conductivity is closely related to the microstructures of the conductors including bulk and interfacial structures. Both ions and electrons can move simultaneously in the materials and are called mixed conductors.

In the 1990s, many ionic conductors were studied and new materials such as BIMEVOX series ($\text{Bi}_4\text{V}_{2-x}\text{M}_x\text{O}_{11}$),

*NEC Corporation.

perovskite-type (ABO_3), and layered rock salt-type (ABO_2) for O^{2-} , H^+ , and Li^+ ionic conductors were developed. New ionic conductors for alkali earth metals (Mg^{+2} , Sr^{+2} , and Ba^{+2}) and trivalent cations (Al^{+3} and Sc^{+3}) were synthesized.

As for the form of ionic conductors, composite and gel increased and research of thin film ionic conductors also increased. New process technology such as mechanical alloying method, rapid quench method, sol-gel method, CVD, and EVD were applied for synthesis of ionic conductors, and the use of fine powders was tried for dense ceramic ionic conductors. Typical ionic conductors are listed in Table I.

Applications of ionic conductors for gas sensor, gas separation, fuel cell, and rechargeable cell are expanding. Oxygen gas sensors of the ZrO_2 system for combustion control of automobile engines are increasing with the enactment of severe regulation for exhaust fumes in many

TABLE I Typical Ionic Conductors

Conductive ions	Typical ionic conductors	Form	Conductivity (S/cm)	Operating temperature (°C)
O^{2-}	$\text{ZrO}_2\text{-Sc}_2\text{O}_3\text{-Al}_2\text{O}_3$	Ceramic	7.6×10^{-2}	800
	$\text{La}_{0.8}\text{Sr}_{0.2}\text{Ga}_{0.8}\text{Mg}_{0.2}\text{O}_3$	Ceramic	2×10^{-2}	800
	$\text{Pb}_{0.8}\text{La}_{0.2}\text{WO}_{4.1}$	Ceramic	4.2×10^{-3}	800
	$\text{BaTh}_{0.9}\text{Gd}_{0.1}\text{O}_{2.95}$	Ceramic	8.7×10^{-2}	550
	$\gamma\text{-Bi}_4\text{V}_{1.7}\text{Sb}_{0.3}\text{O}_{11}$	Ceramic	1×10^{-2}	327
	$\text{Nd}_{0.33}(\text{SiO}_4)_6\text{O}_2$	Crystalline	1.5×10^{-6}	100
F^-	$35\text{InF}_3\text{-}30\text{SnF}_2\text{-}35\text{PbF}_2$	Glass	1×10^{-3}	200
	PbSnF_4	Crystalline	1×10^{-3}	25
H^+	$\text{BaCe}_{0.8}\text{Y}_{0.2}\text{O}_{3-\alpha}$	Ceramic	6×10^{-2}	800
	$\text{W}(\text{PO}_3)_3\text{-PTMO}^a$	Composite	1×10^{-3}	100
	$\text{H}_3\text{PW}_{12}\text{O}_{40}\text{-}29\text{H}_2\text{O}$	Crystalline	2×10^{-1}	25
Li^+	$\text{Li}_{14}\text{Zn}(\text{GeO}_4)_4$	Ceramic	1.3×10^{-1}	300
	$\text{Li}_{2.8}\text{Sc}_{1.8}\text{Zr}_{0.2}(\text{PO}_4)_3$	Ceramic	3×10^{-2}	300
	$\text{Li}_3\text{PO}_4\text{-Li}_2\text{S-Si}_2\text{S}$	Glass	1.8×10^{-3}	25
	$\text{Li}_{0.34}\text{La}_{0.51}\text{TiO}_3$	Ceramic	1×10^{-3}	27
	$\text{Li}_{0.9}\text{NiO}_2$	Ceramic	2.5×10^{-4}	25
	$\text{P}(\text{EO-MEEGE})^b\text{-LiN}(\text{CF}_3\text{SO}_2)_2$	Polymer complex	3×10^{-4}	30
	$(\text{EC-PC})^c\text{/LiClO}_4\text{-PEO}^d$	Gel	2×10^{-3}	20
Na^+	$\text{Na}-\beta''\text{-Al}_2\text{O}_3$	Ceramic	3×10^{-1}	300
	$\text{Na}_3\text{Zr}_2\text{Si}_2\text{PO}_{12}$	Ceramic	5×10^{-2}	200
K^+	$\text{K}_2\text{O-Nd}_2\text{O}_3\text{-SiO}_2$	Ceramic	1.3×10^{-2}	300
Cu^+	$\text{Rb}_4\text{Cu}_{16}\text{I}_{7.2}\text{Cl}_{12.8}$	Ceramic	4×10^{-1}	25
Ag^+	RbAg_4I_5	Crystalline	3×10^{-1}	25
	$82\text{AgI}\text{-}13.5\text{Ag}_2\text{O}\text{-}4.5\text{B}_2\text{O}_3$	Glass	5×10^{-2}	25
	$\text{Mg}_{1.1}\text{Zr}_{3.4}\text{Nb}_{0.6}\text{P}_6\text{O}_{24.4} + 0.4\text{ Zr}_2\text{O}(\text{PO}_4)_2$	Composite	1×10^{-2}	800
Al^{3+}	$\text{Al}_{0.4}\text{In}_{1.6}(\text{WO}_4)_3$	Ceramic	9×10^{-5}	800
Sc^{3+}	$\text{Sc}_{0.33}\text{Zr}_2(\text{PO}_4)_3$	Ceramic	1×10^{-5}	600

^a PTMO, polytetramethyleneoxide.

^b P(EO-MEEGE). poly[ethylene oxide-co-2-(2-methoxyethoxy)ethyl glycidylether].

^c (EC-PC), (ethylene carbonate-propylene carbonate).

^d PEO, polyethylene oxide.

countries. The hydrogen gas sensor of the CaZrO_2 system for meltdown metal is in practical use. Fuel cell using solid oxide is called SOFC. YSZ (yttria-stabilized zirconia) is used at around 1000°C as an ionic conductor for SOFC, and the 5- to 100-kW system has been undergoing testing for a long period of time. ScSZ (Scandia-stabilized zirconia) is promising for improving the capability of electricity generation and the reduction of operation temperatures (650 – 800°C) for long periods of time. $\beta''\text{-Al}_2\text{O}_3$, using Na^+ ion conduction, is used at 300 – 350°C for the NAS (sodium–sulfur) rechargeable battery as electricity storage, and 50- to 500-kW systems are undergoing testing for long periods.

The progress of the lithium rechargeable battery is remarkable with the development of video camcorders, note-type personal computers, and mobile phones because the specific energy of the lithium rechargeable battery is higher than that of the Ni–MH battery. Lithium ionic conductors such as LiCoO_2 , LiNiO_2 , and LiMn_2O_4 are used for cathode materials with the liquid electrolyte. New gel-type lithium ionic conductors have begun to be used instead of the liquid electrolyte.

B. Dielectrics Ceramics

Yukio Sakabe*

1. Ceramic Capacitors

Barium titanate (BaTiO_3) ceramics are still the major dielectrics for advanced ceramics capacitors. Many dielectric of materials are composed of modified dielectrics of BaTiO_3 with the other titanates such as SrTiO_3 , CaTiO_3 , BaTiO_3 ; and zirconate, with BaZrO_3 and CaZrO_3 . A wide variety of dielectric properties have been developed to design the high—performance capacitors. Increasing the dielectric constant, improving the temperature dependence of the characteristics, and designing the highly reliable ceramics were the principal focus for the material design engineers. With the advent of advanced electronic circuits, the demand for surface mountable chip components continues to increase. Therefore, the MLC, which provide a much higher volumetric efficiency than the conventional disk-type (single-layer capacitors), have been manufactured as the major ceramic capacitors since 1988. The structure of this capacitor is shown in Fig. 1. The MLC market has grown by more than 15% every year for the past few years. In 1998, four hundred billion units were consumed in the world.

The most noteworthy work in this capacitor technology in the last decade is the establishment of the manufacturing technology of the MLC with the base metal electrode.

*Murata Manufacturing Co., Ltd.

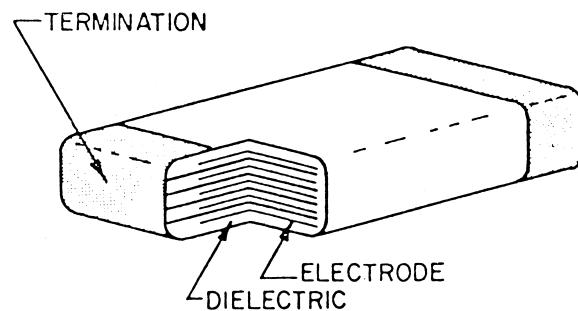


FIGURE 1 Multilayer ceramic capacitors.

Nickel was utilized as an inner electrode with newly developed BaTiO_3 -based dielectrics.

Since 1996, the price of palladium metal for the inner electrode of the capacitors dramatically increased in the market. Replacing the precious metal with a base metal such as nickel is most effective during the manufacturing process. As a result of intensive research on nickel-compatible dielectrics, most of the X7R and Z5U capacitors with high layer counts are now produced with the Ni inner electrode and the Cu termination electrode.

To expand the capacitance range of MLCs to values higher than $10 \mu\text{F}$, process technologies of sheet stacking, more than 300 layers with thin dielectric layers of less than $3 \mu\text{m}$, have been developed. Preparing a very fine particle sized powder, which was synthesized by hydrothermal and hydrolysis methods, has fabricated the ultrathin BaTiO_3 dielectric layer with the Ni electrode. By incorporating $3.0\text{-}\mu\text{m}$ dielectric layers, a $10\text{-}\mu\text{F}$ X7R capacitor in the EIA 1206 size ($3.2 \times 1.6 \times 1.5 \text{ mm}$) has been developed. The large-capacitance MLCs have the potential to replace Ta and Al electrolytic capacitors in the range from 10 to $100 \mu\text{F}$. The major technical advantages over Ta and Al capacitors are higher breakdown voltage, higher reliability, and lower ESR (equivalent series resistance) at high frequencies. The characteristics of the newly developed X7R $10\text{-}\mu\text{F}$ MLCs with the Ni electrode are shown in Table II.

TABLE II Characteristics of X7R $10\text{-}\mu\text{F}$ Multilayer Ceramic Capacitors (MLC) with Ni Electrodes

Chip size	EIA 1206 ($3.2 \times 1.6 \times 1.5 \text{ mm}$)
Thickness of dielectric layer	$3.0 \mu\text{m}$
Thickness of inner electrode	$1.3 \mu\text{m}$
Layer count	350
DC rated voltage	6.3 V
Capacitance (1 kHz, 1 V rms)	$10.5 \mu\text{F}$
Dissipation factor (1 kHz, 1 V rms)	3.5%
CR product (25°C)	$5300 \Omega \times \text{F}$
Breakdown voltage (DC)	>170 V

TABLE III Microwave Properties of Dielectric Resonator Materials

Materials	ϵ_r	$Q \times f$ (GHz)	τ_f (ppm/ $^{\circ}$ C)
MgTiO ₃ –CaTiO ₃ system	21	60,000	0–6
Ba(Mg,Ta)O ₃ system	24	200,000–350,000	0–6
Ba(Zn,Ta)O ₃ system	30	100,000–180,000	0–6
Ba ₂ Ti ₉ O ₂₀	38	50,000	4
(Zr,Sn)TiO ₄	38	60,000	0–6
CaTiO ₃ –NdAlO ₃	43	45,000	0–6
Ba(Sm,Nd) ₂ Ti ₄ O ₁₂	80	100,000	0–6
(Ba,Pb)Nd ₂ Ti ₄ O ₁₂	92	5,000	0–6
(Ba,Pb)(Nd,Bi) ₂ Ti ₄ O ₁₂	110	2,500	0–6

2. Dielectric Resonator

Dielectric resonators have gained a position as key elements in microwave components for size reduction in microwave filters and as frequency-stabilizing elements in oscillator circuits. Dielectric resonators reduce the physical size of resonant systems because the electromagnetic wavelength is shortened in dielectrics to $1/\sqrt{\epsilon_r}$ of its value in free space, where ϵ_r is the dielectric constant of the resonator.

The required properties for dielectric resonator materials are (1) high dielectric constant, (2) low dielectric loss tangent $\tan \delta$, and (3) low temperature coefficient of resonant frequency τ_f . Many kinds of dielectric resonator materials have been developed since the 1970s. **Table III** shows the dielectric properties of some materials that are now commercially available. In the table, the quality factor Q is reciprocal of dielectric loss tangent; $Q = 1/\tan \delta$. As the $\tan \delta$ is proportional to frequency for ionic paraelectric materials, the product of Q and frequency is the value inherent to each material. Some materials have as high Q values as the copper cavity, and some have the temperature coefficient as stable as the invar cavity. The material with lower ϵ_r generally has a higher Q value.

It is known that the complex permittivity of dielectric resonator materials follows the dielectric dispersion equation, which is expressed as the superposition of electronic and ionic polarization.

$$\dot{\epsilon}(\omega) - \epsilon(\infty) = \frac{\omega_T^2(\epsilon(0) - \epsilon(\infty))}{\omega_T^2 - \omega^2 - j\gamma\omega}, \quad (1)$$

where ω_T and γ are the resonance frequency and damping constant of the infrared active lattice vibration modes, and $\epsilon(\infty)$ is the permittivity due to the electronic polarization.

As the condition $\omega^2 \ll \omega_T^2$ is reasonably adopted at microwave and millimeter-wave frequencies, the following equations are derived from Eq. (1).

$$\epsilon'(\omega) - \epsilon(\infty) = \epsilon'(0) - \epsilon(\infty). \quad (2)$$

$$\tan \delta = \frac{\epsilon''(\omega)}{\epsilon'(\omega)} = \frac{\gamma}{\omega_T^2} \omega. \quad (3)$$

These equations show that the dielectric constant is independent of frequency and the dielectric loss tangent increases proportionately to frequency at microwave frequency. **Figure 2** shows the frequency dependence of $\tan \delta$ from 1 to 35 GHz for some materials. As seen in Eq. (3), $\tan \delta$ increases proportionately to frequency.

Figure 3 shows the temperature dependence of resonance frequency and $\tan \delta$. Superior temperature stability of resonant frequency is obtained by selecting the composition of each material. The temperature coefficient of resonant frequency τ_f is defined by the following equation:

$$\tau_f = -\frac{1}{2}\tau_\epsilon - \alpha, \quad (4)$$

where τ_ϵ is the temperature coefficient of the dielectric constant and α is the linear thermal expansion coefficient of the dielectric specimen.

The intrinsic origin of the damping constant γ in Eq. (3) is the anharmonic terms in the crystal's potential energy by which the phonons of infrared active modes decay into the thermal phonons. As the phonon energy follows the Planck distribution, γ , $\tan \delta$ increases proportionately to temperature when the abnormality such as phase transition does not occur in the observed temperature range. **Figure 4** shows the temperature dependence of $\tan \delta$ from 30 to 300 K.

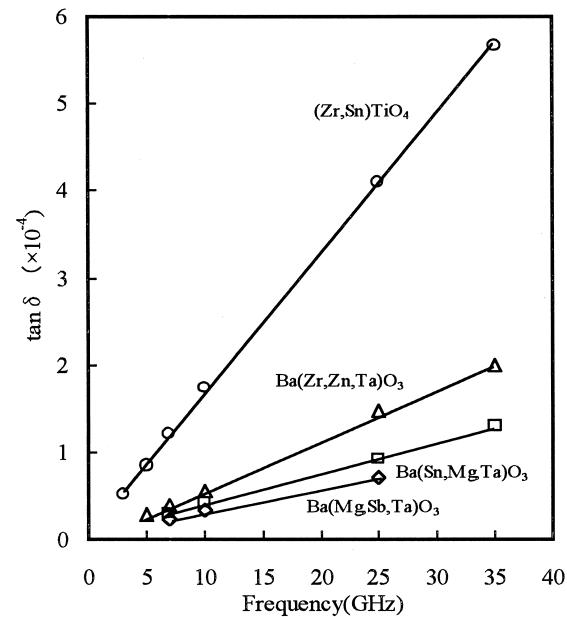
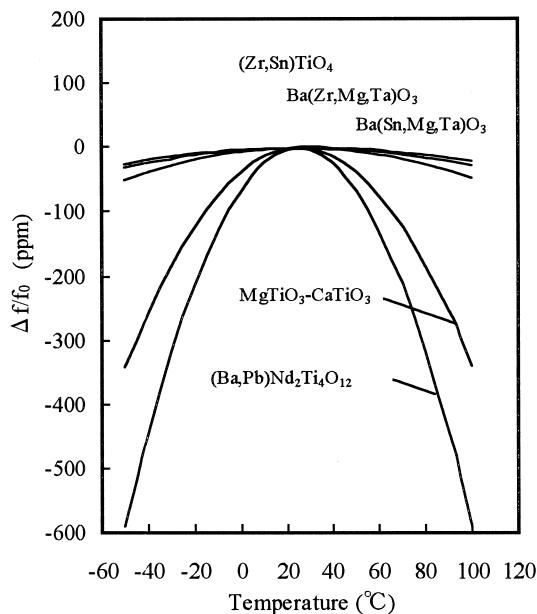
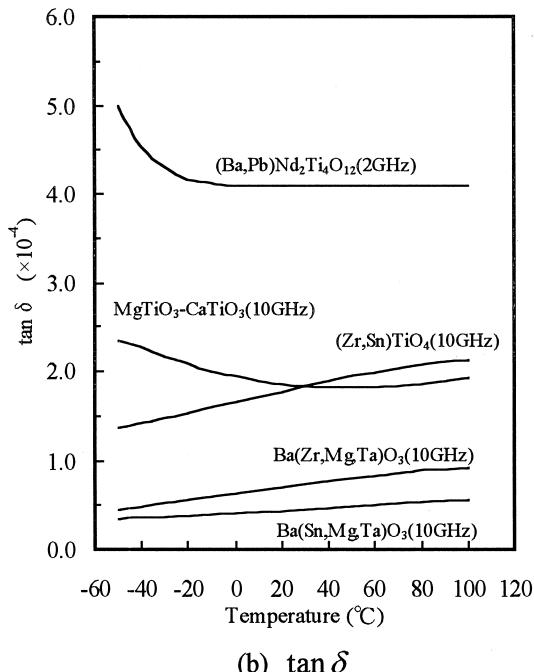


FIGURE 2 Frequency dependence of $\tan \delta$ from 1 to 35 GHz.

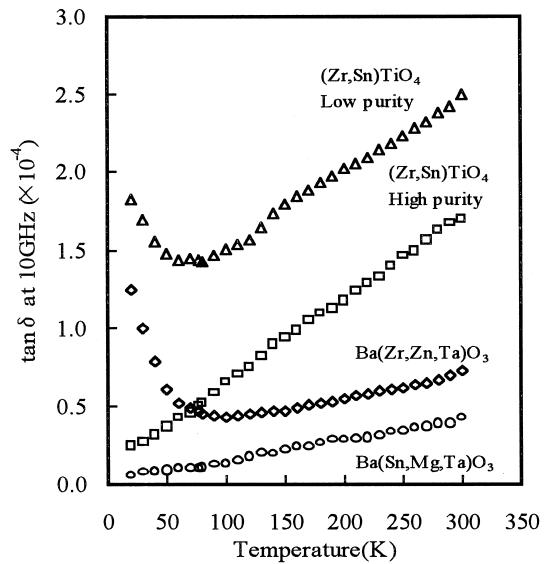


(a) resonance frequency

FIGURE 3 Temperature dependence of resonant frequency and $\tan \delta$.

The main applications for dielectric resonator materials are the antenna duplexers of cellular mobile phones, filters of cellular base stations, oscillators of DBS-TV converter, and filters at millimeter-wave frequencies.

The materials of the $(\text{Pb},\text{Ba})(\text{Bi},\text{Nd})_2\text{Ti}_4\text{O}_{12}$ system with high ε_r are popularly used for antenna duplexer of

FIGURE 4 Temperature dependence of $\tan \delta$ from 30 to 300 K.

cellular mobile phones. The materials $(\text{Zr},\text{Sn})\text{TiO}_4$ and $\text{CaTiO}_3-\text{NdAlO}_3$ with high Q and high ε_r are used for filters of cellular base stations. The complex perovskite materials, $\text{Ba}(\text{Zn},\text{Ta})\text{O}_3$ and $\text{Ba}(\text{Mg},\text{Ta})\text{O}_3$ with very high Q values, are used for applications higher than 10 GHz.

C. Semiconductive Ceramics

Yukio Sakabe*

1. PTC Thermistors

Barium titanate (BaTiO_3) is a ferroelectric material with a high dielectric constant and a high insulation resistance. Therefore, it has been widely used in the electrical industry for ceramic capacitors since its discovery in 1943. Small amounts of rare earth metal oxide such as Sm_2O_3 , CeO_2 , Y_2O_3 and La_2O_3 convert the insulating BaTiO_3 ceramic into a semiconductor with resistivity in the range of 10 to $10^6 \Omega \cdot \text{cm}$.

This behavior has been harmful because dielectric materials require high insulation resistance and low dielectric loss. In 1955, the unusual temperature dependence of resistance above the Curie temperature of semiconductive BaTiO_3 ceramics was discovered.

This behavior of resistance was a very exciting discovery not only for practical application of thermistors but also for fundamental research of conduction mechanisms.

This semiconductor called the PTC (positive temperature coefficient) thermistor undergoes a drastic resistance

*Murata Manufacturing Co., Ltd.

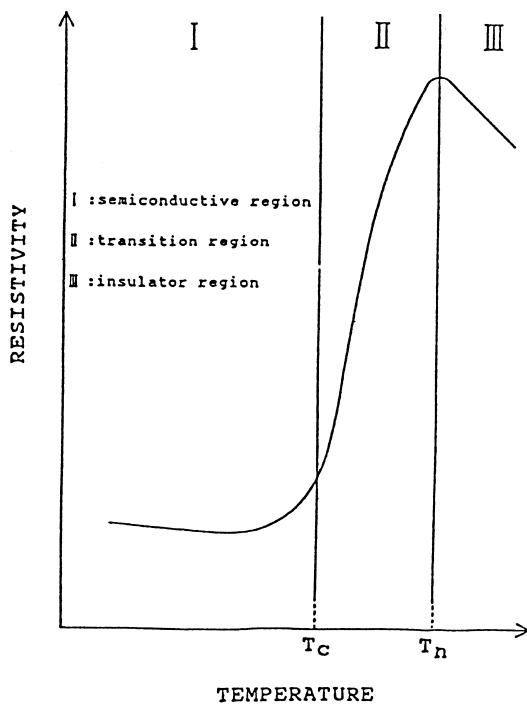


FIGURE 5 Characterized temperature region or PTC ceramics.

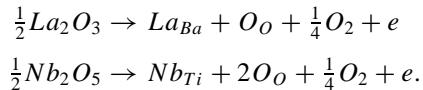
increase at the Curie temperature (T_c), and its resistance increases up to the temperature where the resistance reaches its maximum value. The characterized temperature is divided into three regions (I, II, and III in Fig. 5) according to the resistance behavior.

The ferroelectric phase transition and a microstructure of the ceramic play an important part in these PTC characteristics. In temperature region I ($T < T_c$), its temperature coefficient is usually negative. This is expressed as

$$R = R_0 \exp B_1 \left(\frac{1}{T} - \frac{1}{T_0} \right),$$

where R_0 is the resistance at reference temperature (T_0), and B_1 is the thermistor constant.

To produce semiconductive BaTiO_3 , a small amount of rare earth metal ions (e.g., Y^{3+} or La^{3+}) are substituted at the Ba^{2+} site, or Nb^{5+} and Ta^{5+} ions are substituted at the Ti^{4+} site. This reaction is given in the following equations, using the defect notation of Kröger and Vink.



In temperature region II ($T_c < T < T_n$), resistance across the grain boundary increases exponentially with increasing temperatures. This is expressed as

$$R = R_0 \exp A(T - T_0),$$

where R_0 is the resistance at the reference temperature (T_0), and A is the temperature coefficient of resistance ($A > 0$).

The increase of resistance corresponds to the decrease of spontaneous polarization (P_s) of BaTiO_3 due to the phase transition from the ferroelectric tetragonal phase to the paraelectric cubic phase.

The strong internal field caused by P_s depresses the potential barrier height. The gradual decrease of P_s and the dielectric constant, expressed by the Curie–Weis law, cause the potential barrier height to recover, which results in an increase of resistance in region II.

In region III ($T > T_n$) resistance decreases again from maximum resistance with

$$R = R_0 \exp B_2 \left(\frac{1}{T} - \frac{1}{T_0} \right),$$

where B_2 is called thermistor constant.

The transition temperature T_c of BaTiO_3 can be lowered or elevated from its original value (120°C) by substitution of Sr^{2+} or Pb^{2+} ions at the Ba^{2+} site, respectively. Figure 6 shows the typical resistivity versus temperature curves of the PTC thermistor.

a. Manufacturing process. Reagent grade BaCO_3 , TiO_3 and donor dopant (e.g., Y_2O_3 or La_2O_3) are used

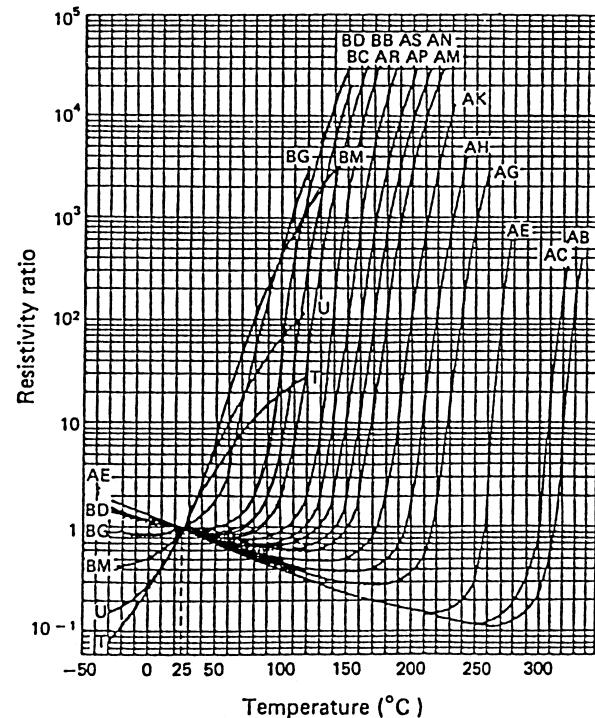


FIGURE 6 Resistance–temperature characteristics of PTC ceramics.

as starting materials. The manufacturing processes are almost the same as those in the electronic ceramics industry.

To control the resistance and the temperature coefficient within the exact range, much attention is paid to the firing temperature and ensuing cooling rate. Base metal electrodes such as Ni, Zn, and Al provide ohmic contact with PTC thermistor, which in n-type semiconductor ceramics.

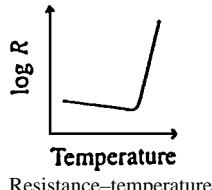
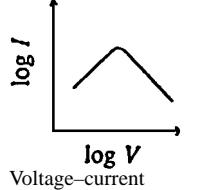
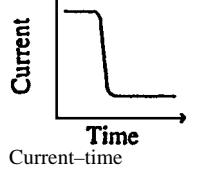
b. Applications. The PTC thermistor has three useful functions:

1. Drastic changing of resistance at characterized temperature T_c .
2. Self-controlling of temperature.
3. Limiting current through the body.

These functions have been applied for many electronic devices as shown in [Table IV](#).

c. Future development. The lower resistance of PTC at room temperature is required to limit the current in low-voltage circuits. A miniaturized 0.1Ω PTC thermistor will soon be manufactured. To meet the requirements of surface mounting technology (SMT), a chip-type PTC thermistor will be developed and applied to hybrid circuits.

TABLE IV Functions and Applications of PTC Ceramics

Function	Applications
 Resistance-temperature	Temperature sensing Temperature indicating Temperature measurement Overheat protection Power transistor, Transformer
 Voltage-current	Electric limiter Protection against overcurrent Constant temperature heating Electrical heater, Curing iron
 Current-time	Motor starter Degaussing of color TV Delay time switch

2. NTC Thermistor

The negative temperature coefficient (NTC) thermistor is semiconductive materials whose resistance decreases with increasing temperature. They usually consist of ceramics composed of metal oxides such as MnO_2 , NiO , Co_2O_3 , and Fe_2O_3 .

The resistance can vary with different compositions and sizes of ceramic elements between 10 and $10^6 \Omega$ at room temperature with temperature coefficients between -2 and $-6.5\%/\text{ }^\circ\text{C}$.

The thermistor elements are commercially available in a wide variety of sizes ranging from beads of 0.2 mm in diameter through disks and washers with diameters between 3 and 25 mm , to rods up to 12 mm in diameter and 40 mm long. Bead types are encapsulated in solid glass. Disk devices are coated with an insulating lacquer or epoxy resin.

The chip-type NTC thermistor has become popular because of suitability for SMT. The inner electrode type, especially, offers high-resistance accuracy and high reliability.

[Figure 7](#) shows a comparison of resistance change with temperature between various types of thermistors. Temperature dependence of resistance is given by

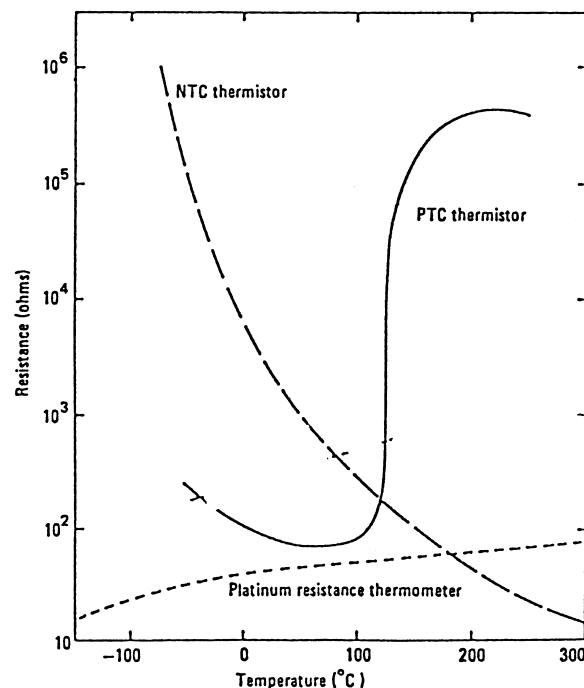


FIGURE 7 Temperature dependence of different types of thermistors in contrast to a platinum resistor.

$$R = A \exp(B/T) \quad (1)$$

$$B = E/k, \quad (2)$$

where A is a constant, B a thermistor constant, E the activation energy, and k Boltzmann's constant. The temperature coefficient of resistance, α , is also defined as

$$\alpha = \frac{1}{R} \frac{dR}{dT} \quad (\%/\text{°C}) \quad (3)$$

From Eqs. (1) and (3), rewritten with B

$$\alpha = -B/T^2. \quad (4)$$

a. Processing. The starting materials for manufacture of NTC ceramics are MnO_2 , NiO , Co_2O_3 , Fe_2O_3 , CuO , and ZnO . These are wet-mixed with ball mill, then calcined at 700 to 900°C. The calcined powder is mixed again with organic binder. Dried powder is pressed into disks, then sintered at 1000 to 1300°C.

The crystal structure is a spinel. The final stage of ceramic element is formation of the electrode. Usually, Ag paste is printed onto both surfaces, then fired again at 500 through 900°C.

The chip-type NTC thermistor is manufactured by forming into sheets and laminating with the electrode.

b. Applications. One of the most important fields of NTC thermistor applications is temperature sensing, either for the purpose of temperature measurement, temperature control, and thermal alarms, or for the compensation of changes in circuit resistance brought about by variations in the ambient temperature. All of these applications are based on the resistance-temperature characteristics of these thermistors. They could also be widely applied as components, for example, as an electric delay, a surge suppressor, liquid or gas flow sensor, or liquid level sensor.

c. Trends. The trends in modern NTC ceramics have resulted in thermistor development in three directions. The first trend is the improvement of device stability. The second is the high accuracy of resistance and B -constant. The third is the miniaturization of the chip-type NTC thermistor.

3. Ceramic Varistor

Metal oxide varistor is a ceramic semiconductor device having highly nonlinear current-voltage characteristics expressed as

$$\begin{aligned} I &= (V/C)^\alpha \\ \alpha &= \frac{dI/I}{dV/V} = \frac{d \log I}{d \log V} \\ &= \frac{\log I_2 - \log I_1}{\log V_2 - \log V_1}, \end{aligned}$$

where α is the nonlinear exponent, C the constant corresponding to the resistance, and V_1 and V_2 are the voltages at the currents of I_1 and I_2 , respectively. The nonlinear resistance C is conveniently given by V_c , that is a voltage per unit length (V/mm) when 1 mA/cm² of current flows through the body. Thus the ceramic varistor is characterized by the nonlinear exponent α and varistor voltage V_c .

Silicon carbide ceramics have been used for varistors for many years, but their nonlinear exponent α is about 6, which is not enough to protect the electronic circuits against surges. Back-to-back Zener diodes were preferably used as a protector device against surges because of their high α value of more than 100. Their varistor voltage V_c and energy capabilities are rather small, hence they are not suitable in the power circuits.

Zinc oxide-based ceramic varistors were developed in 1970. They are characterized with a highly nonlinear voltage-current relationship and large energy capabilities. Their α value is in the range of 40 to 50, and the V_c adjustable to values in the range from 50 to 250 V/mm.

Strontium titanate-based varistors were developed in 1980. The feature of these varistors is their larger electrostatic capacitance compared with the other types. The SrTiO_3 ceramics are essentially dielectrics with a dielectric constant of 320, which is much higher than that of ZnO . Firing in a reducing atmosphere converts the SrTiO_3 to semiconductive ceramics. Figure 8 shows the voltage-current characteristics of the various varistors.

a. Structure of the ZnO varistors. These varistors are produced by the general method of the manufacturing of electroceramics. The mixture of ZnO powder and the small addition of Bi_2O_3 and some other oxides is sintered at temperatures in the range of 1100 to 1300°C.

Their electrical properties are very sensitive to the ceramic microstructure. The structure, shown in Fig. 9, results from a liquid-phase sintering due to the presence of molten Bi_2O_3 and other dopants.

b. Electrical properties. The ZnO grains ($\sim 10 \mu\text{m}$ in diameter) are n-type semiconductors. The intergranular traps are formed in this grain boundary due to the presence of Bi_2O_3 , and transition metal oxides, which cause the double Schottky barrier.

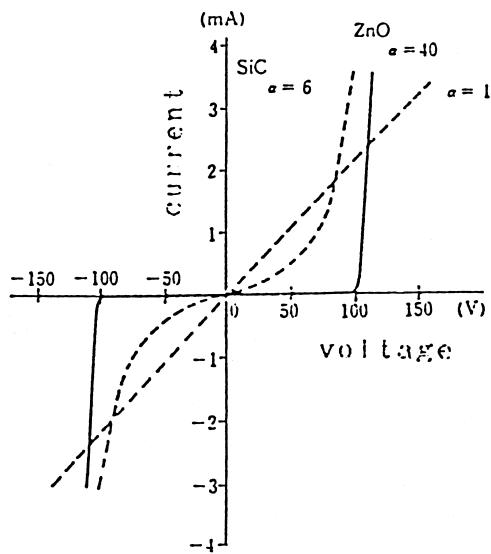


FIGURE 8 Current versus voltage characteristics of ceramic varistors.

c. Applications. Metal oxide varistors are mainly used in circuits for protection against inductive surges, very short spike noise, or power surges. They are always connected across the power line in parallel with the load to be protected. This is indicated schematically in Fig. 10. A varistor should be chosen that has a varistor voltage V_c slightly higher than the signal voltage applied to the load to be protected. The varistor is insulating in normal operation where the applied voltage is lower than V_c . If a transient pulse, whose voltage is higher than V_c , is incident, the varistor current rapidly increases, resulting in a conducting shunt path for the incident pulse.

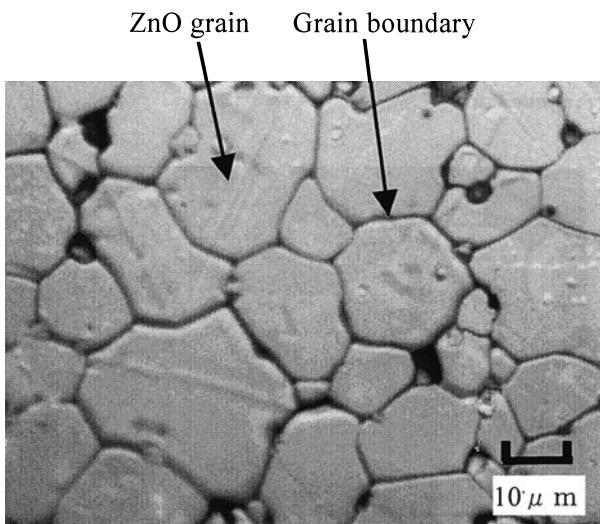


FIGURE 9 SEM micrograph of ZnO-based ceramic varistor.

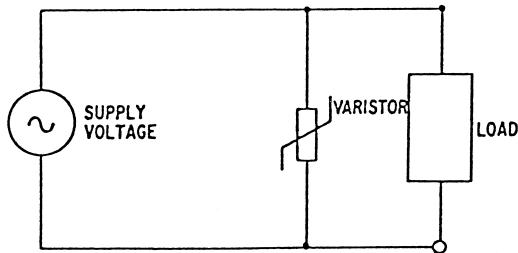


FIGURE 10 Typical application of ZnO varistor as a transient protective device.

D. Piezoelectric Composites

Y. Yamashita*

1. Introduction

Since the discovery of ferroelectricity of polycrystalline ceramics (BaTiO_3) during the early 1940s, there has been a continuous succession of new piezoelectric materials, such as $\text{Pb}(\text{Zr}_{1-x}, \text{Ti}_x)\text{O}_3$ (PZT) (Jaffe *et al.*, 1955) and $\text{Pb}(\text{B}'\text{B}'')\text{O}_3$, ($\text{B}' = \text{Zn, Mg, In and Sc}$ and $\text{B}'' = \text{Nb, Mo, Ta, and W}$), so called the relaxor materials, in ceramics and single crystals form, which has led to a significant commercial application in the world (Kuwata *et al.*, 1982, and references herein).

The potential piezoelectric properties of the perovskite (ABO_3) structure of PZT ceramics, near the morphotropic phase boundary (MPB) separating the rhombohedral and tetragonal phases, were first clarified by Jaffe *et al.* in 1955. After the discovery of PZT, numerous $\text{Pb}(\text{B}'\text{B}'')\text{O}_3$ and $\text{Pb}(\text{B}'\text{B}'')\text{O}_3\text{-PbTiO}_3$ materials similar to PZT were also investigated by Russian scientists in the 1950s–1960s (Galasso, 1969). Important materials systems reported in the early days were $\text{Pb}(\text{Zn}_{1/3}\text{Nb}_{2/3})\text{O}_3\text{-PbTiO}_3$ (PZN-PT or PZNT) (Kuwata *et al.*, 1982), $\text{Pb}(\text{Ni}_{1/3}\text{Nb}_{2/3})\text{O}_3\text{-PbTiO}_3\text{-PbZrO}_3$ (PNN-PT-PZ) and $\text{Pb}(\text{Mg}_{1/3}\text{Nb}_{2/3})\text{O}_3\text{-PbTiO}_3\text{-PbZrO}_3$ (PMN-PT-PZ) (Ouchi *et al.*, 1965), for piezoelectric applications. Recently, one of the present authors reported that the 1 mol% PbNb_2O_6 -doped $\text{Pb}(\text{Sc}_{1/2}\text{Nb}_{1/2})\text{O}_3\text{-PbTiO}_3$ ceramic material near the MPB prepared by modern processing exhibits excellent piezoelectric properties, including $k_p > 71\%$, $k_t > 54\%$, $k_{33} > 77\%$, $d_{33} > 500 \text{ pC/N}$, and $T_c = 250^\circ\text{C}$ (Yamashita, 1994a,b). These are superior piezoelectric properties compared to conventional PZT ceramics, which were previously believed to be the best binary piezoelectric ceramic materials. The section of piezoelectric ceramics introduces a recent progress of the $\text{Pb}(\text{B}'\text{B}'')\text{O}_3\text{-PbTiO}_3$ ceramics and single crystals materials. Their piezoelectric properties and composition

*Toshiba Corporation, Kawasaki 210-8582 Japan.

relations are summarized. The present situation and future outlook for these materials are also presented.

2. Pb(B'B'')O₃-PbTiO₃ Ceramic Materials

Table V shows the Curie temperature T_c , maximum dielectric constant (Kmax), crystal structure, ferroelectric (F) or antiferroelectric (AF) phase, PbTiO₃ mol% at the MPB, and T_c at the MPB for various relaxor materials, along with their abbreviations. Figure 11 shows the MPB (I) in PZT and the MPB (II) in Pb(B'B'')O₃-PbTiO₃ and a comparison of the two. The location of the MPB (II) varies when different B' or B'' ions are selected. As shown in Table V, the MPB (II) compositions were paid little attention as regards practical piezoelectric application until the 1980s, due to their inferior piezoelectric properties, especially their low electromechanical coupling factors and higher raw materials cost compared to those of PZT ceramics.

Table VI indicates the five high-coupling factors (k_s), perovskite binary material systems, and two ternary material systems with planar coupling factors $k_p > 60\%$ and/

or $k_{33} > 70\%$ reported recently. It is interesting to note that these seven materials have a number of similarities. Table VII lists the similarities among these high k_s materials. The first similarity is the rhombohedral and tetragonal MPB. Aside from PMN which has a pseudo-cubic structure, PZ, PZN, PST, and PSN have a rhombohedral structure. This rhombohedral structure has 8 possible polarization directions. The tetragonal structure has 6 possible polarization directions. Therefore, a MPB consisting of both phases has 14 possible polarization directions which may contribute to a large k_s . The second similarity is large remanent polarization. The seven high k_s material systems, PZT 53/47, PZNT 91/9, PMNT 67/33, PSNT 58/42, PSTT 55/45, PSMNT29/34/37 (Yamashita *et al.* (1996b)), and PSZNT29/40/31 (Yamashita *et al.*, 1998), have a high value of $Pr > 25 \mu\text{C}/\text{cm}^2$. It is well known that the polarization is roughly proportional to the magnitude of the atomic displacement and this contributes to large dielectric constant. Therefore, a large Pr value is a basic requirement for a high k_s material. The third similarity is

TABLE V Various Relaxor Materials and Their Properties

Relaxor	Abbrev.	T_c (°C)	Kmax	Crys. Structure	F or AF ^a	Ti @ MPB Ti (mol%)	T_c @ MPB (°C)
Pb(B ²⁺ _{1/3} B ⁵⁺ _{2/3})O ₃							
Pb(Cd _{1/3} Nb _{2/3})O ₃	PCdN	270	8,000	PC	F	28	380
Pb(Zn _{1/3} Nb _{2/3})O ₃	PZN	140	22,000	R	F	9–10	180
Pb(Mg _{1/3} Nb _{2/3})O ₃	PMN	−10	18,000	PC	F	32–34	155
Pb(Ni _{1/3} Nb _{2/3})O ₃	PNM	−120	4,000	PC	F	30–35	130
Pb(Mn _{1/3} Nb _{2/3})O ₃	PMnN	−120	4,000	PC	F	30–35	130
Pb(Co _{1/3} Nb _{2/3})O ₃	PCoN	−98	6,000	M	F	33	250
Pb(Mg _{1/3} Ta _{2/3})O ₃	PMgT	−98	7,000	PC	F	30?	100
Pb(B ³⁺ _{1/2} B ⁵⁺ _{1/2})O ₃							
Pb(Yb _{1/2} Nb _{1/2})O ₃	PYbN	280	150	M	AF	50	360
Pb(Lu _{1/2} Nb _{1/2})O ₃	PLuN	260	350	M	AF	41	350
Pb(In _{1/2} Nb _{1/2})O ₃	PIN	90	550	M	F	37	320
Pb(Sc _{1/2} Nb _{1/2})O ₃	PSN	90	38,000	R	F	42	260
Pb(Fe _{1/2} Nb _{1/2})O ₃	PFN	112	12,000	R	F	7?	140
Pb(Sc _{1/2} Ta _{1/2})O ₃	PST	26	28,000	R	F	45	205
Pb(B ²⁺ _{1/2} B ⁶⁺ _{1/2})O ₃							
Pb(Cd _{1/2} W _{1/2})O ₃	PCW	400	400	M	AF	?	
Pb(Mg _{1/2} W _{1/2})O ₃	PMgW	39	300	O	AF	55	60
Pb(Co _{1/2} W _{1/2})O ₃	PCoW	32	240	O	AF	45	310
Pb(B ³⁺ _{2/3} B ⁶⁺ _{1/3})O ₃							
Pb(Fe _{2/3} W _{1/3})O ₃	PFW	−75	9,000	C	F		
Others							
*PbZrO ₃	PZ	240	3,000	R	AF	47	360
(Pb,La)(Zr,Ti)O ₃	PLZT	<350	30,000	R,T	F,AF	30–45	<350
PbTiO ₃	PT	490	9,000	T	F		

^a R, rhombohedral; T, tetragonal; PC, pseudocubic; M, monoclinic; O, orthorhombic; F, ferroelectrics; AF, antiferroelectrics.

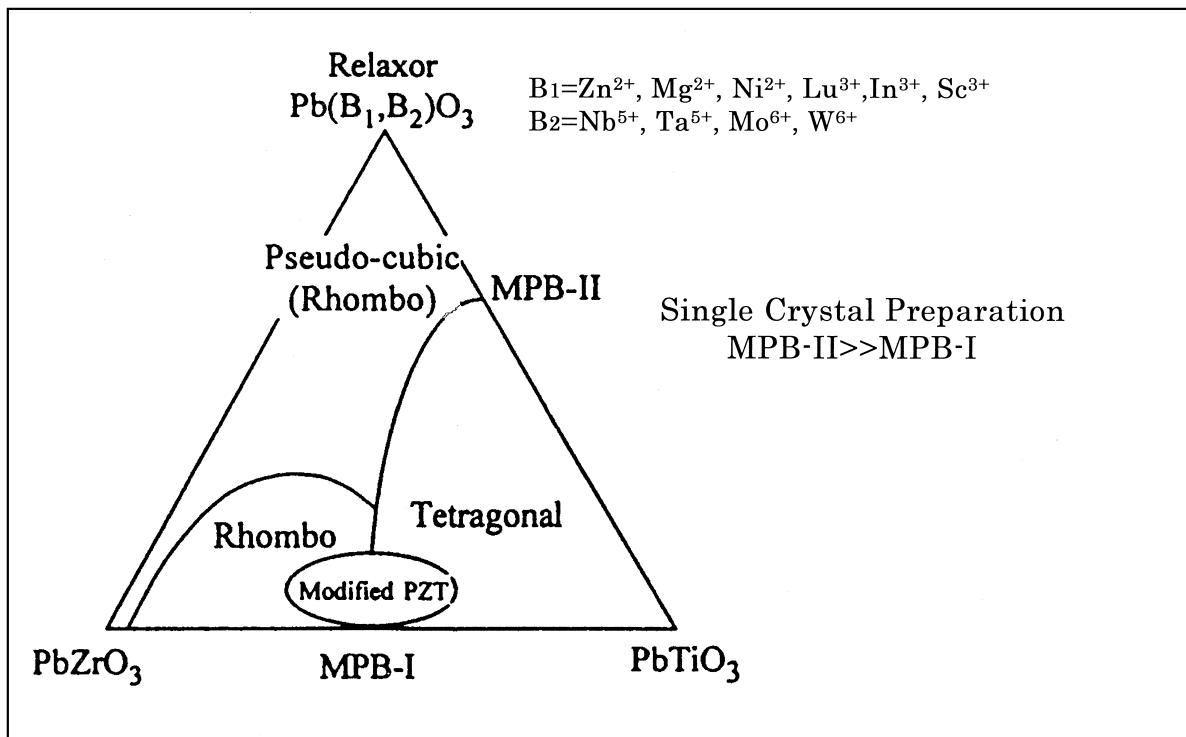


FIGURE 11 Morphotropic phase boundary MPB (I) in PZT and MPB (II) in $\text{Pb}(\text{B}'\text{B}'')\text{O}_3$ -PT system and their comparison.

that all the large k_s materials have similar B' and B'' ionic sizes. In the ABO_3 perovskite structure, the A-site ion is 12 coordination and the B-site ion is 6 coordination. Ionic sizes of these values were reported by [Shannon \(1976\)](#).

The PZT contains Ti with a small size (60.5 pm) and Zr with a large ionic size (74 pm). When the small ionic size elements Sn (69 pm) or Hf (71 pm) are placed in the Zr sites in PZT, the coupling factors decrease. It is interesting to note that ionic sizes of all large k_s B' ions, Zr, Zn, Mg, and Sc, are similar and in the range 70–74 pm. In relaxor-PT systems, relaxors containing smaller B' -size elements such as Ni (69 pm), Fe (65 pm), and Co (65 pm) do not exhibit large k_s values at the MPB. Relaxors with larger-ionic-size B' element, Y (89 pm) are unlikely to have a perovskite structure. Recently, Yamamoto *et al.* reported on the piezoelectric properties of the $\text{Pb}(\text{Yb}_{1/2}\text{Nb}_{1/2})\text{O}_3-\text{PbTiO}_3$ ceramic system. The maximum electromechanical coupling factor k_p is 56% near the MPB composition of PYbNT 50/50 ([Yamamoto and Ohashi, 1995](#)). Niobium has the largest B'' ionic size, 64 pm, of all $\text{Pb}(\text{B}',\text{B}'')\text{O}_3$ relaxors. There have been few reports of high k_s values in relaxor-PT systems consisting of B'' atoms of smaller size, such as W (60 pm) or Mo (61 pm). Therefore, Nb may be the best B'' element for achieving high k_s in relaxor-PT systems. The fourth similarity is the average ionic size at the MPB. The average ionic sizes of all seven large k_s materials are in the relatively narrow

range of 64.7–66.8 pm. The last similarity is the light average molecular mass of B-site ions. High k_s materials have a smaller molecular mass (MM) of average B-site MM, less than 80 ([Yamashita *et al.*, 2000](#)). Complex Perovskite of the $x\text{Pb}(\text{Sc}_{1/2}\text{Nb}_{1/2})\text{O}_3-y\text{Pb}(\text{Mg}_{1/3}\text{Nb}_{2/3})\text{O}_3-z\text{PbTiO}_3$ (PSMNT 100x/100y/100z) ternary ceramic materials near the MPB. The MPB traces an almost linear region between the PSNT 58//42 and PMNT 68/32 MPB compositions of the binary system. The maximum piezoelectric constant, $d_{33} = 680 \text{ pC/N}$, was found at PSMNT 29/33/38, where $\varepsilon_{33}/\varepsilon_0 = 3800$, and the electromechanical coupling factors $k_p = 70\%$, $k_{31} = 43\%$, and $k_{33} = 76\%$, and $T_c = 215^\circ\text{C}$ were obtained ([Yamashita, 1996b](#)). [Yamamoto and Yamashita \(1996\)](#) reported very large electromechanical coupling factors, $k_p = 76\%$, $k_{33} = 79\%$, $k_{31} = 46\%$, and $k_t = 56\%$, of hot-pressed PSNT58/42 ceramics. The outstanding properties of the ceramics also had a small grain size of 1.0μ . Such a small grain size with large electromechanical coupling is very difficult to obtain in PZT materials. The k_p value of the PSMNT is the highest among all the perovskite ceramics ever reported and better than those for PZT ceramics.

3. $\text{Pb}(\text{B}',\text{B}'')\text{O}_3-\text{PbTiO}_3$ Single Crystals

The solid solution of PZT is an oxygen-octahedra ferroelectric with a perovskite structure. Many studies have been made using ceramic specimens prepared by

TABLE VI Large Electromechanical Coupling Factors k_p , k_{33} , and Piezoelectric Constant d_{33} of Binary and Ternary Material Systems

No.	Materials	Feature	k_p (%)	k_{33} (%)	d_{33} (pC/N)	T_c (°C) ^a	References
1	PZT53/47	Ceramics	0.52	0.67	220	370	Jaffe <i>et al.</i> , 1955
	PZT53/47 + Nb	Ceramics	0.67	0.76	400	360	
	PZT50/50	Single Crystal	?	?	?	360	Tsuzuki <i>et al.</i> , 1973
2	PZNT88/12	Ceramics	(0.5)	0.48	235	222	
	PZNT 91/9	Single Crystal(001)		0.92	1500	180	Harada <i>et al.</i> , 1998; Kobayashi <i>et al.</i> , 1997; Kuwata <i>et al.</i> , 1982
3	PMNT67/33	Ceramics	0.63	0.73	690	160	Choi <i>et al.</i> , 1989
	PMNT70/30	Single Crystal(001)		0.94	1500	145	Shroud <i>et al.</i> , 1990
4	PSTT55/45	Ceramics	0.61	0.73	655	205	
	PSTT55/45	Single crystal	?	(0.85)	(>1000)		
5	PSNT 58/42	Ceramics	0.71	0.77	540	260	Yamashita, 1994a,b
		Ceramics(Hot press)	0.76	0.79	620	250	Yamamoto and Yamashita, 1996
		Single crystal	?	(>0.94)	(>1000)		Yamashita and Shimanuki, 1996a
6	PInNT 64/36	Ceramics	0.58	0.70	440	280	Alberta and Bhalla, 1998
	PInNT 72/28	Single crystal		(0.85)	700	250	Yasuda <i>et al.</i> , 2000
7	PSMNT	Ceramics	0.72	0.78	640	215	Yamashita <i>et al.</i> , 1996b
	29/34/37	Single crystal	?	(>0.94)	(>1500)	215	Yamashita <i>et al.</i> , 2000

^a Estimated value.

solid-state reactions, but the electrical properties of single crystals of PZT have yet to be studied sufficiently (Tsuzuki *et al.*, 1973). This may be principally due to the difficulty of growing large enough crystals to enable measurement of their electrical properties. On the contrary, making $\text{Pb}(\text{B}'\text{B}'')\text{O}_3$, PbTiO_3 , and the solid solution of $\text{Pb}(\text{B}'\text{B}'')\text{O}_3\text{-PbTiO}_3$ in the form of binary single crystals is relatively easy. Many reports of complex perovskites, such as PMN, PZN, PNN, PFN, and PSN have been made (Alberta and Bhalla, 1998; Galasso, 1969; Harada *et al.*, 1998; Kobayashi *et al.*, 1997; Kuwata *et al.*, 1982; Park and Shroud, 1997a,b; Shimanuki *et al.*, 1998; Shroud *et al.*, 1990; Yamashita and Shimanuki, 1996a; Yasuda *et al.*, 2000).

The most important report relating to $\text{Pb}(\text{B}'\text{B}'')\text{O}_3\text{-PbTiO}_3$ binary single crystals considers PZNT91/9, written by Kuwata *et al.* (1982). Using the single crystals, they measured for the first time a large difference in electrical properties between the [001] and [111] planes. A sample poled along the [001] axis revealed an unexpectedly large piezoelectric constant ($d_{33} > 1500$ pC/N) and elec-

tromechanical coupling factor ($k_{33} = 0.92$) at room temperature in the rhombohedral phase. Shroud *et al.* (1990) studied the piezoelectric properties of PMN-PT single crystals near the MPB. The PMNT 70/30 single crystal exhibits almost the same value of $d_{33} > 1500$ pC/N as the [001] plane. Figure 12 represents the piezoelectric material tendencies of various piezoelectric materials. In the early 1950s, barium titanate was the only material known to offer a large d_{33} (>200 pC/N). After the discovery of PZT in 1955, d_{33} values reached 400–800 pC/N. However, in the last 20 years, no big change in d_{33} has been seen in piezoelectric ceramics. Single crystals of PZNT 92/8-90/10 and PMNT 70/30-67/33 have a large d_{33} (>2000 pC/N), which has never before been obtained with ceramic specimens. This is one reason why an intensive search for new high d_{33} or k_{33} single crystals is important. Many kinds of high-performance single crystals, such as PZNT, PSNT, PSMNT, and PInNT have been reported by many scientists in the last 5 years. (Harada *et al.*, 1998; Hosono *et al.*, 2000; Kobayashi *et al.*, 1997; Shimanuki *et al.*, 1998; Yasuda *et al.*, 2000). Large sizes, more than 40 mm in diameter, of PZNT and PMNT single crystals have been developed (Harada *et al.*, 1998).

TABLE VII Five Similarities of High k_s Materials

1	Rhombohedral-tetragonal MPB
2	Large Polarization $\text{Pr} > 25 \mu\text{C}/\text{cm}^2$
3	Proper Range of B' ions (70–74 pm), Disorder?
4	Appropriate average ionic size in MPB (64.7–66.8 pm)
5	Lighter molecular mass of B-site ions in MPB (MM < 80)

4. Application of New $\text{Pb}(\text{B}'\text{B}'')\text{O}_3\text{-PbTiO}_3$ Materials

Piezoelectric application of $\text{Pb}(\text{B}'\text{B}'')\text{O}_3\text{-PbTiO}_3$ materials has a few commercial products which many research

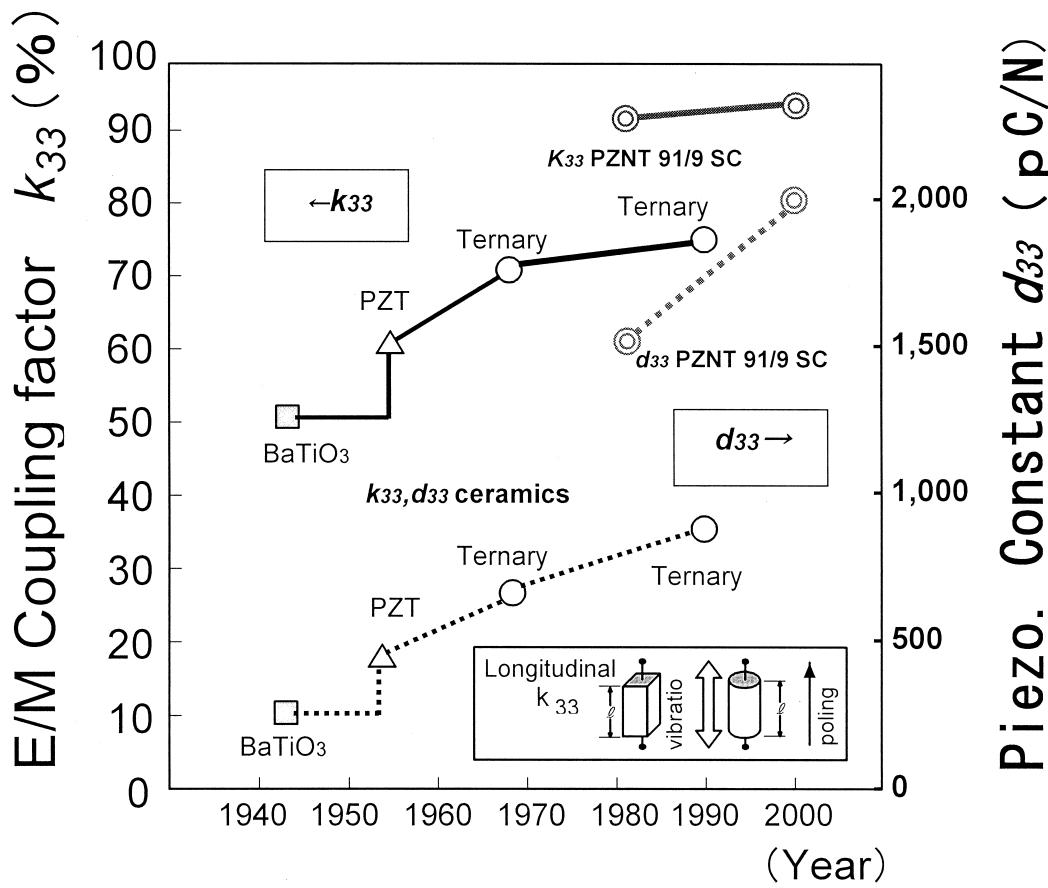


FIGURE 12 Piezoelectric characteristics trend of various piezoelectric materials.

works have reported over the last 20 years. The biggest problem of $\text{Pb}(\text{B}'\text{B}'')\text{O}_3\text{-PbTiO}_3$ materials compared to conventional PZT or PLZT is its raw material cost. Powder cost of commercial PZT is \$20–40/kg. In comparison, $\text{Pb}(\text{B}'\text{B}'')\text{O}_3\text{-PbTiO}_3$ materials powder cost is \$50–600 kg depending on the composition. Therefore, high-frequency applications are suitable because they require a relatively small amount of powder. For example, the raw material cost of a 2- to 5-MHz medical transducer element (25.4 mm in diameter and 0.4 mm in thickness, about 1.6 g) is less than \$1.00 even using the highest raw material powder of the PSNT 58/42. Therefore, the most preferable area of the first application of $\text{Pb}(\text{B}'\text{B}'')\text{O}_3\text{-PbTiO}_3$ materials will be medical echo ultrasonic transducers where high performance, rather than raw cost, is the top priority. The advantages of an extremely large electromechanical coupling factor ($k_{33} > 84\%$) and a low acoustic impedance ($Z_{33} < 23 \times 10^6 \text{ kg/m}^2 \text{ s}$) of the PZNT single-crystal materials for medical ultrasonic transducers are characterized as better acoustic impedance matching, broader band width, and better resolution (Park and Shrout, 1997a,b; Saitoh *et al.*, 1999). The second promising area is actuator applications. The PZNT and PMNT single crystals show

extraordinary large strain versus voltage. The large strain with minimum hysteresis is quite attractive for positioners and several undersea applications (Park and Shrout, 1997a).

Another promising area is thin film application. Since some relaxor thin film has a large dielectric constant compared to SiO_2 , Ta_2O_5 , or BaSrTiO_3 (BST), researchers have also rediscovered the utility of relaxor thin film materials as high-dielectric-constant capacitors, which opens up new possibilities for manufacturing planar, very high-density random-access memories (DRAMs) (Takeshima *et al.*, 1995).

Bibliography for Section I.D can be found at the end of this article.

E. Piezoelectric Ceramics

Yukio Sakabe*

Piezoelectric ceramics are known for piezoelectric and inverse piezoelectric effects. The piezoelectric effect produces an electric displacement subjected to mechanical

*Murata Manufacturing Co., Ltd.

stress. On the other hand, the inverse piezoelectric effect makes a crystal to produce a strain by applying an electric field. Ceramics have a multi-crystalline structure. Since each crystal in the ceramics orients at random after firing, no piezoelectric effects are produced in the as-fired ceramics. However, when ceramics are ferroelectric, piezoelectricity can be added after firing. When a higher electric field than the coercive field of a ferroelectric ceramic is applied, a polarization reversal occurs, and the value of remnant polarization of the ceramics is not zero, and piezoelectricity is produced. This treatment is called the poling treatment.

1. Piezoelectric Ceramic Materials

Of piezoelectric materials, Rochelle salt and quartz have long been known as a single-crystal piezoelectric. However, these materials have had a relatively limited application range mainly because of the poor crystal stability of Rochelle salt and the limited degree of freedom in the characteristics of quartz. Later barium titanate (BaTiO_3), the first discovered piezoelectric ceramic, was introduced for applications in ultrasonic transducers, mainly for fish finders. Lead zirconate titanate system ($\text{PbZr}_{1-x}\text{Ti}_x\text{O}_3$) ceramics were discovered in the late 1950s. This ceramic system is usually abbreviated as PZT. PZT has electromechanical transformation efficiency and stability far superior to existing materials.

Another popular piezoelectric ceramic is lead titanate, which is one of the most thermally stable ferroelectric materials because its Curie temperature is around 500°C . However, some compositional modification is necessary for improving its sinterability.

2. Applications of Piezoelectric Ceramics

a. Ceramic filters. Ceramic filters are devices for selecting specific electrical wave frequencies. Today, most AM and FM radio receivers and sound-video circuits of TV sets use ceramic filters as intermediate frequency (IF) filters.

The quality of this electronic communication equipment is determined by the selectivity and stability of the IF filter. Ceramic filters satisfy these requirements and make the IF filter circuit adjustment free. Piezoelectric ceramics act as both a mechanical resonator and an electromechanical transducer that responds to a specific frequency.

Piezoelectric ceramics have many resonant vibration modes corresponding to their size. Ceramic filters are widely used as the LF, MF, HF, and VHF band filters.

i. Low-frequency band filter. LF (30–300 kHz) band filters normally use the lengthwise vibration mode of a rectangular plate of PZT ceramics.

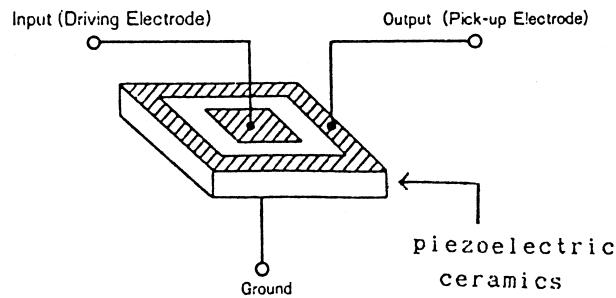


FIGURE 13 MF band filter with split electrode on the square piezoelectric ceramic.

The fundamental resonance frequency of the rectangular ceramic with its length of l , is

$$f = \frac{1}{2l} \sqrt{\frac{E}{\rho}},$$

where E is Young's modulus and ρ is the density of the ceramics.

ii. Medium-frequency band filter. MF (300 kHz–3 MHz) band filters normally use a radial vibration mode. IF filters (455 kHz) for the AM radio are the main products at these frequencies.

The three terminal band pass filter has a split electrode on a thin ceramic disk or square plate. The resonance frequency f is

$$f = \frac{\varphi}{\pi D} \sqrt{\frac{E}{\rho(1 - \sigma^E)}} \quad (\text{disk-type})$$

$$f = \frac{1}{2l} \sqrt{\frac{E}{\rho(1 - \sigma^E)}} \quad (\text{square-plate-type}),$$

where σ^E is Poisson's ratio and $\varphi = (1 - \sigma^E)J_1(\varphi) = \varphi J_0(\varphi)$.

Figure 13 is shows the schematic diagram of this ceramic filter, and its equivalent circuit is shown in Fig. 14. Here, C_{o1} and C_{o2} are dumped capacitance, and C_{o12} is stray capacitance between input and output terminal. The resonance circuit is composed of the equivalent components, that is, capacitance C_1 , inductance L_1 , resistance R_1 , and transformer.

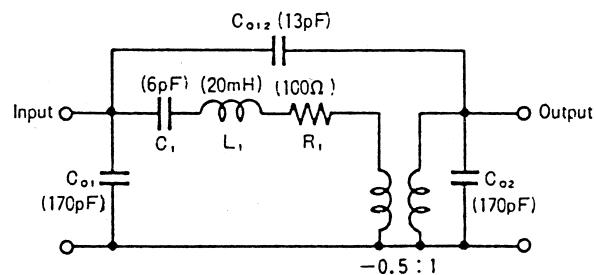


FIGURE 14 Equivalent circuit of the MF band ceramic filter.

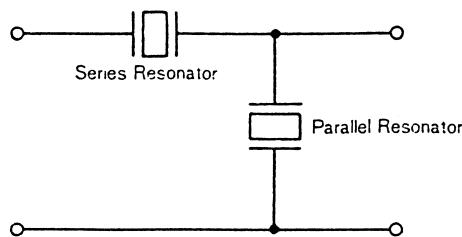


FIGURE 15 Single segment of the ladder-type MF ceramic filter.

Ladder-type ceramic filters consist of many simple segments electrically connected in series. Each segment is composed of two piezoelectric ceramic resonators connected in series and parallel as shown in Fig. 15. The resonance frequency of the parallel resonator is adjusted to the anti-resonance frequency of the series resonator to provide the four terminal bandpass filter. The frequency bandwidth is determined by the frequency difference between resonance f_r and anti-resonance f_a of the resonators. The attenuation is determined by capacitance ratio between the series and the parallel resonators and determined by the number of the segments.

Recently the piezoelectric longitudinal effect of a multilayered piezoelectric ceramic rod has been utilized as the resonator of the ladder filter. This multilayered resonator dramatically decreased the dimension of MF filter.

iii. High-frequency bandfilter. The HF band (3 MHz–30 MHz) includes sound for TV IF (4.5 MHz), FM IF (10.7 MHz), and citizen band (27 MHz). Therefore, many ceramic filter are used in this band region. The resonance frequency of the thickness mode is given by

$$f = \frac{1}{2t} \sqrt{\frac{1}{c_{33}^D \cdot \rho}},$$

where t is the thickness of the plate, c_{33}^D the elastic stiffness constant, and ρ the density.

Trapped thickness vibration mode is based on the theory of *energy trapped vibration*. Vibration energy is localized at only the electrode area. Today, most of the ceramic filters for this frequency band use this trapped thickness mode because the spurious response is eliminated and selectivity is enhanced.

Figure 16 shows a cross section of a multicoupling energy trapped mode ceramic filter. Two vibration modes are excited in a resonator with a split electrode; one is symmetrical and the other is anti-symmetrical. By the coupling of both modes, a band-pass filter is obtained.

b. Ceramic resonators. Ceramic resonators, which use piezoelectric ceramic plates as mechanical vibrators, are elements of highly stable oscillation circuits. Ceramic resonators utilize the intrinsic mechanical resonance of

the piezoelectric ceramic plate (which is determined by its dimensions), the elasticity, and the vibrating mode of the plate.

Oscillation circuits with ceramic resonators have been used mainly as clock-generators for various microprocessors, which are represented by one-chip microcomputers.

The features of ceramic resonators are as follows.

1. High stability of oscillating frequency. (The temperature coefficient of the oscillating frequency of the circuit in which the ceramic resonator is used is less than $10^{-5}/^\circ\text{C}$.)
2. The value of the mechanical quality factor is between that of quartz resonators and LC or CR resonators.
3. Oscillation rises quickly in approximately $\frac{1}{10}$ – $\frac{1}{100}$ the time of that in quartz resonators.
4. Ceramic resonators are small and light, approximately half the size of quartz resonators.
5. Using ceramic resonators in transistors or various ICs allows for nano-adjustment oscillation circuits. (There is no need for potentiometers and trimmer capacitors.)

Because of these excellent features, ceramic resonators are widely used (e.g., for TVs, VCRs, CD-players, DVD-players hard disc drives (HDDs), automobile electronic components, telephones, cellular phones, cameras, camcorders, various remote control systems, and so forth.)

The vibrating mode of the piezoelectric vibrator (of the ceramic resonator) is selected according to the oscillating frequency. For instance, if oscillating frequency is desired to be less than 1 MHz and more than 100 kHz, the radial vibration mode is suitable. The relation of the vibrating mode and the oscillating frequency is shown in Fig. 17.

The impedance-frequency characteristic and the phase-frequency characteristic of ceramic resonators are shown Fig. 18. The frequency at which impedance of a ceramic resonator takes minimum value is called the resonance

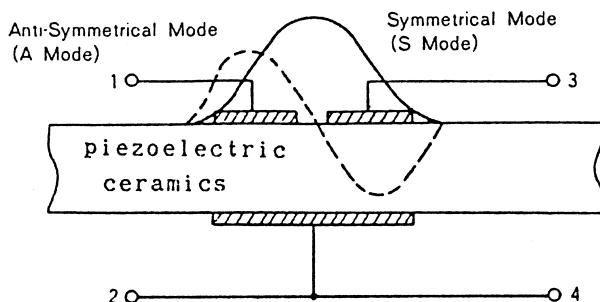


FIGURE 16 Multicoupling mode filter with energy trapped vibration mode.

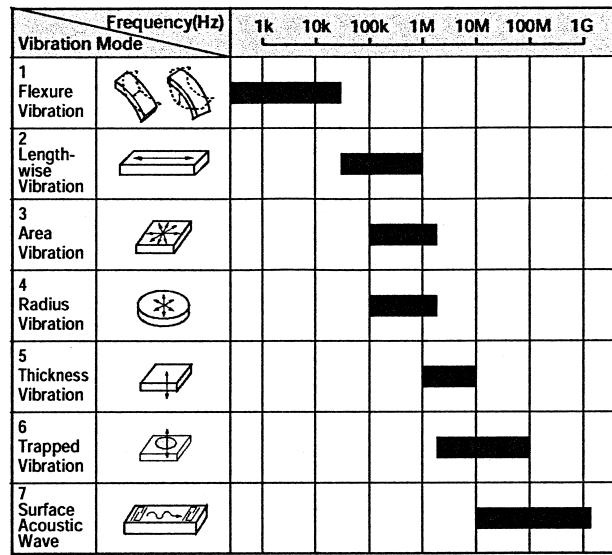


FIGURE 17 Relation of vibrating mode and frequency range.

frequency f_r , and the frequency at which impedance of the resonator takes maximum value is called the antiresonance frequency f_a . From Fig. 18 it can be seen that ceramic resonators are inductive (like coils) in their frequency range between f_r and f_a and they become capacitive (like capacitors) in the remaining frequency range. This means that a ceramic resonator, in a frequency range near the resonance frequency, can be represented by an equivalent circuit with a coil L_1 and a resistor R_1 , and two

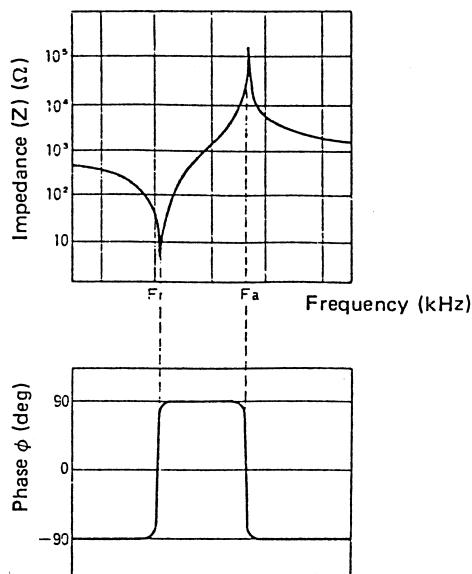


FIGURE 18 Impedance frequency and phase frequency characteristics of a ceramic resonator.

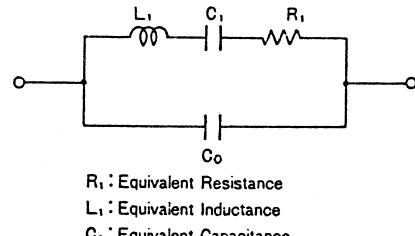


FIGURE 19 Electrical equivalent circuit for a ceramic resonator.

capacitors C_0 and D_1 as shown in Fig. 19. The frequencies are expressed as follows:

$$\begin{aligned} f_r &= \frac{1}{2}\pi \cdot \sqrt{L_1 \cdot C_1} \\ f_a &= \frac{1}{2}\pi \cdot \sqrt{L_1 \cdot C_1 \cdot C_0 / (C_1 + C_0)} \\ &= f_r \cdot \sqrt{1 + C_1 / C_0} \end{aligned}$$

In the frequency range between f_r and f_a a ceramic resonator is considered as an inductor L_e (H) with internal loss R_e (Ω). In the case of ceramic resonators, the value of inductance of the equivalent inductor varies approximately from 0 to 10^{-2} H. Since this operation takes place selectively in the frequency range between f_r and f_a (of the whole frequency range), an oscillation circuit can be composed of a ceramic resonator.

Common oscillation circuits with coils and capacitors are Colpitts circuits or Hartley circuits (see Fig. 20). By replacing coils in Fig. 20 with ceramic resonators, oscillating conditions are satisfied selectively in the frequency range between f_r and f_a .

The effective temperature stability of inductance of a ceramic resonator is much higher than that of a coil. So highly stable oscillation circuits, in which the temperature coefficient of the oscillating frequency is less than $10^{-5}/^\circ\text{C}$, are composed of ceramic resonators. Oscillating conditions are found by analogy for LC oscillation circuits.

c. Ceramic buzzer. The piezoelectric buzzer produces sound using the vibration characteristics of piezoelectric ceramic plates to produce acoustical vibration.

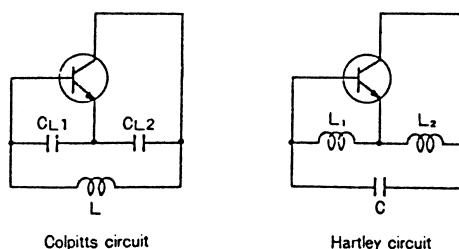


FIGURE 20 Basic configuration of LC oscillation circuit.

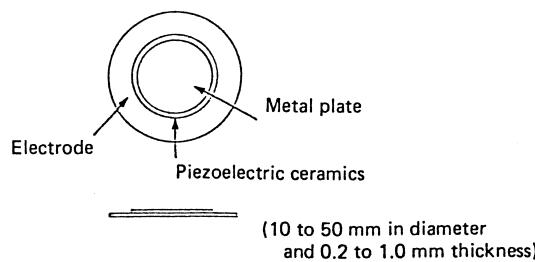


FIGURE 21 Piezoelectric buzzer element.

The vibration is connected to mechanical resonance, (determined by the dimensions) of the structures.

The buzzer elements is constructed of a thin metal plate adhered to thin plate of piezoelectric ceramics which have electrodes on both sides. The dimensions of a typical piezoelectric buzzer are shown in Fig. 21.

Electrodes are formed on both sides and then the thin piezoelectric ceramics (e.g., ferroelectric PZT material) are polarized in the direction of thickness. When an external signal is applied to piezoelectric ceramics in the direction of polarization, the ceramics distort both in the direction of the polarization and in the direction perpendicular to the polarization. The latter distortion is the fundamental movement of a piezoelectric buzzer. Because a piezoelectric buzzer element is stuck on a metal plate which dose not distort itself, as shown in Fig. 22, a bending vibration occurs when an electric signal is applied. It vibrates the surrounded air and produces sound.

In a free vibration, however, the sound pressure produced by the piezoelectric buzzer is small because its acoustic impedance does not match that of air.

Some steps can be taken to increase sound pressure. To make a complete piezoelectric buzzer, therefore, the element is assembled in a resonance box. The typical example of a supported buzzer element and resonance box is shown in Fig. 23. The dimensions of the mechanical resonance system are very important for sound produced by a piezoelectric buzzer element and a resonance box in order to obtain high power at resonance frequency.

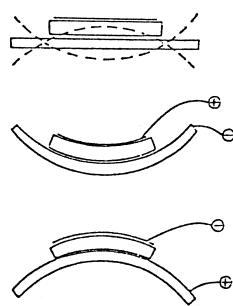


FIGURE 22 Vibration in bending mode.

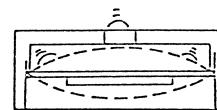


FIGURE 23 An example of supporting buzzer element (edge support).

Resonance frequency of a piezoelectric buzzer is determined physically by quality and dimension of the piezoelectric ceramics and the metal plate given by

$$f = \frac{\alpha^2 h}{4\sqrt{3}\pi a^2} \frac{Y}{\rho(1 - \sigma^2)},$$

where f is the resonance frequency of the buzzer element, a diameter, h the thickness, Y Young's modulus, ρ the density, σ Poisson's ratio, and α a constant.

Until now the reproducibility of piezoelectric buzzers has been poor in mass production, particularly in terms of the characteristics and economics.

Recently, however, these problems have been solved; namely, the technique used in mass producing thin plates of ceramics and the technique of processing ceramics are highly advanced, and it has come to be possible to produce inexpensive piezoelectric buzzers with fine characteristics.

Compared with the electromagnetic buzzer, the advantages of the piezoelectric buzzer are as follows.

1. Low energy consumption
2. Simple structure; enables thin devices
3. The excellence in reliability and life

As a result, piezoelectric buzzers are employed in watches, various kinds of ringers, pagers, and many alarm systems.

d. Piezoelectric actuator. A piezoelectric actuator produces displacement and force by the inverse piezoelectric effect. Most piezoelectric actuators are classified into two types. One is a direct extensional type piezoelectric element; the other is a flexure type.

The typical direct extensional type actuator is shown in Fig. 24. It is composed of many thin piezoelectric ceramic plates electrically bonded in parallel and mechanically connected in series so that their displacements are accumulated. As a result, the actuator is largely displaced in the direction of length. The actuator was first applied to the impact printer head element, and has broadened its application fields such as the fine displacement apparatus, and so forth. It is also a strong candidate of the magnetic head actuator for a high-density HDD.

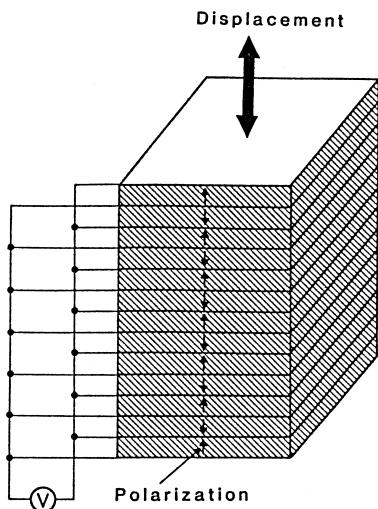


FIGURE 24 Thickness extensional type piezoelectric actuator.

Typical flexure-type actuators are bender bimorphs and unimorphs. The unimorph is a two-plate assembly with only one piezoelectric; a piezoelectric buzzer is an example. The bender bimorph (Fig. 25) is composed of two, thin, bonded piezoelectric ceramics plates. By applying voltage to the ceramics plates, the actuator vibrates in bending mode. It is used as the actuator for automatic focusing systems of camcorders, and so forth.

The actuator for automatic focusing systems (Fig. 26) consists of two annular piezoelectric ceramic plates for driving and an annular metal plate for a support. The piezoelectric ceramic plates are bonded to both sides of a metal plate by means of a bonding agent. An optical system lens or an image sensor is fitted in the central bore in the metal plate of the actuator.

The actuator is supported firmly at the outer peripheral edge of the metal plate near the outer peripheral portion of the piezoelectric ceramics plate. By applying the driving voltage to a piezoelectric ceramic plate, the lens or the image sensor is finely displaced in the optical axis direction. The automatic focusing is performed by analyzing the change in the signal intensity on the image sensor during the vibration.

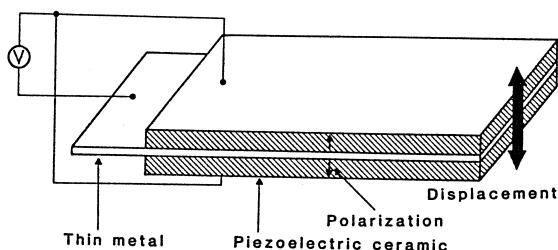


FIGURE 25 Bender bimorph type piezoelectric actuator.

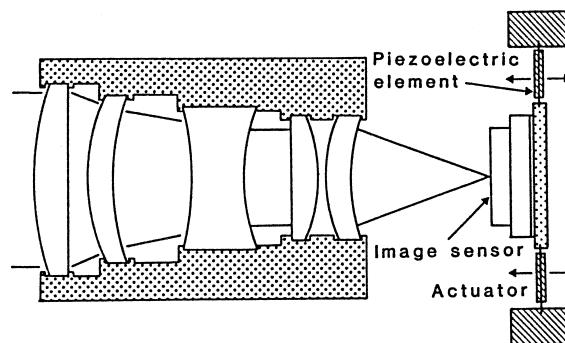


FIGURE 26 Automatic focusing system.

Recently piezoelectric actuators have been widely used in printing heads of ink-jet printers. Both the direct extensional type and the flexure type are now in use. The direct extensional type adopts the multilayered structure, and produces a large displacement and large force with a high speed. Thick film forming technique is utilized in the flexure-type piezoelectric actuator. A piezoelectric ceramic paste is screen-printed on a thin ceramic portion above the cavity. After a heat treatment, a piezoelectric flexural actuator is prepared. When an electric field is applied to the piezoelectric film, a cavity volume changes by the flexural distortion, and ink drops are spit out.

e. Piezoelectric transformer. Piezoelectric transformers have been used in the backlight inverter for the liquid crystal displays (LCDs). Their efficiency reaches up to 90%, and energy consumption of a piezoelectric transformer is much less than a conventional electromagnetic transformer, whose efficiency is typically around 70%.

The output voltage of a piezoelectric transformer changes depending on load resistance. This characteristic is suitable for Cold cathode fluorescent lamps (CCFL) which need a high voltage over 1 kV, at the moment when the switch is turned on, and needs 200–300 V during lighting.

Additionally, since thinner inverters are possible with piezoelectric transformers, they are increasing their application fields, such as mobile computers or handy VCR cameras or palm-top displays (electronic memo-pads). The Rosen-type or modified Rosen-type piezoelectric transformer (Fig. 27) is mainly used in these application fields.

f. Piezoelectric sensor. Piezoelectric sensors are classified into two types. One consists of two parts, which are an ultrasound generating part and a receiving (detecting) part. The other has only the detecting part, and it directly utilizes the piezoelectric effect. The ultrasonic sensors which are used in back sonar systems of automobiles,

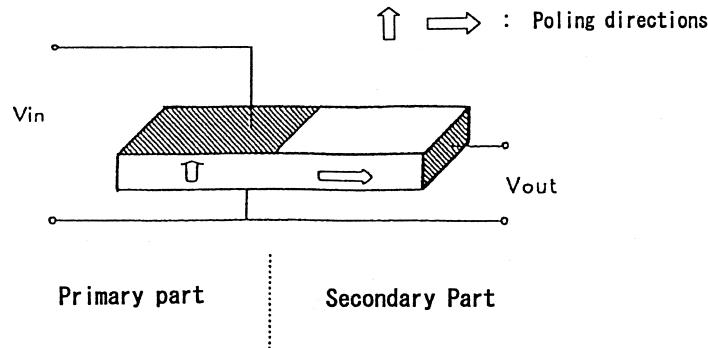


FIGURE 27 Rosen type piezoelectric transformer.

gas flow meters, medical equipment, and so forth, belong to the former. A piezoelectric gyroscope also belongs to this kind of sensor. It is an angular velocity sensor, which is used in navigation systems for automobiles, and a vibration compensation system of a camcorder. It has both a wave generating part and a wave receiving part. In this case, the generated ultrasonic wave propagates in the gyroscope element itself, and the detecting part detects the modified wave by Coriolis force of the rotation. Various kinds of piezoelectric gyroscope structures have been proposed. Examples of these are shown in Fig. 28.

The piezoelectric sensor, which consists of only a detecting part, is often called a piezoelectric type. It directly utilizes a piezoelectric (right) effect. Typical sensors of

this type are a knocking sensor of an automobile engine and a shock sensor. Figure 29 shows a newly developed piezoelectric shock sensor for HDD.

F. Low Firing Multilayer Ceramic Substrates

Kazuaki Utsumi and Hideo Takamizawa*

Low firing multilayer ceramic substrates are ceramic substrates made with glass-ceramic materials which can be sintered at around 900°C.

Alumina green sheets have been developed which show promise for use in high-density packaging substrates.

*NEC Corporation.

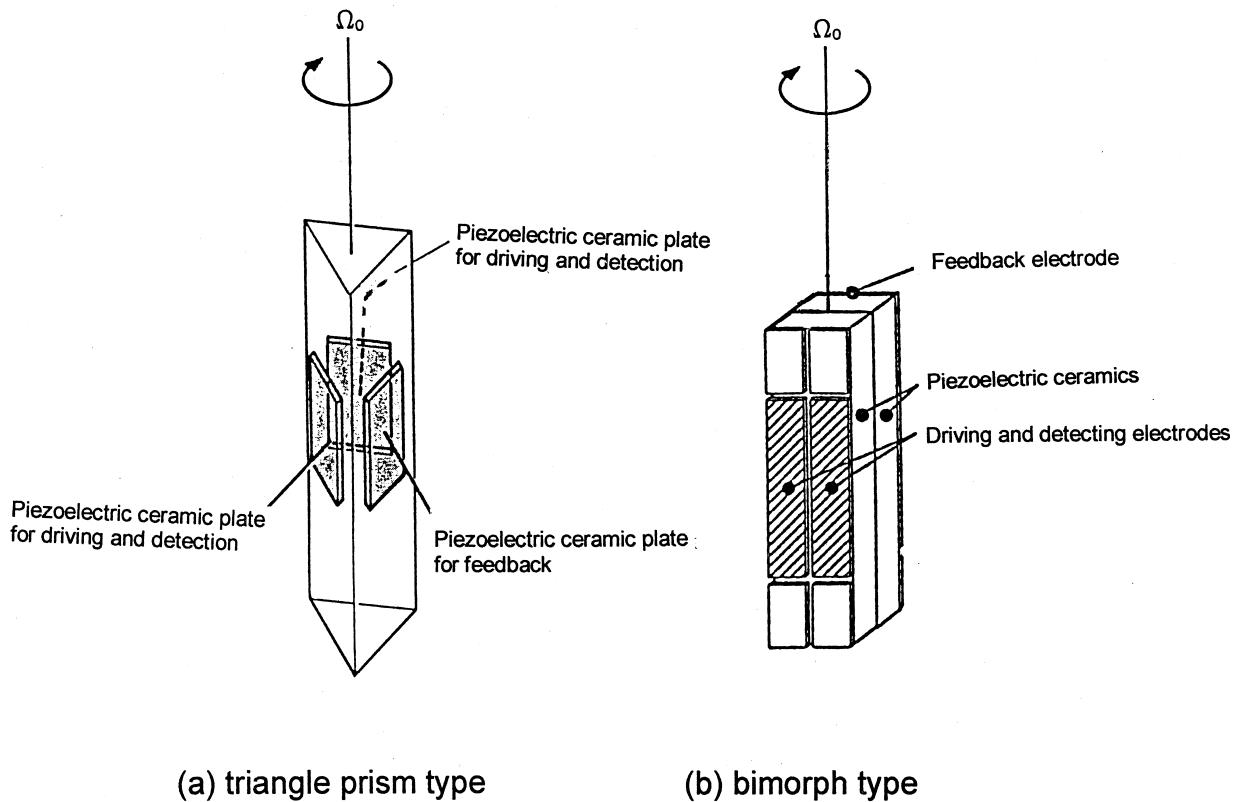


FIGURE 28 Piezoelectric gyroscope structures.

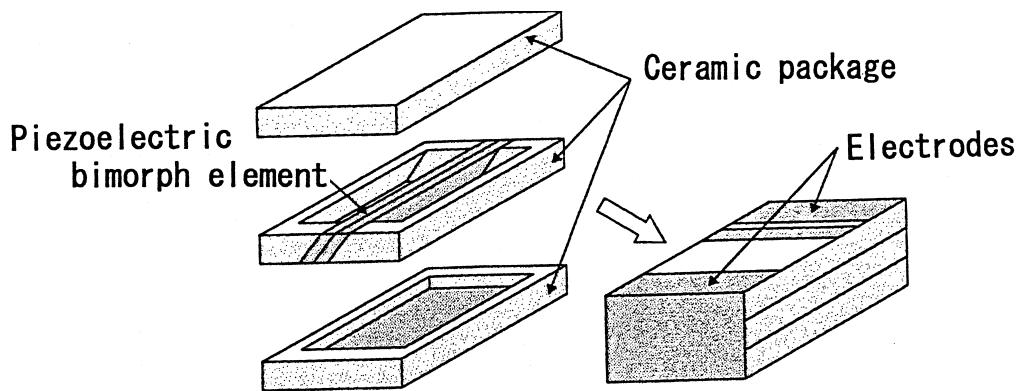


FIGURE 29 Piezoelectric shock sensor.

Since this material sinters at high temperatures (1500°C or above), it is necessary to use molybdenum or tungsten as the conductor and to sinter in a reducing atmosphere (dry hydrogen) in order to prevent oxidation of these metals. The thermal expansion coefficient of alumina is about $70 \times 10^{-7}/^{\circ}\text{C}$, much larger than that for silicon semiconductor chips. The ability to reduce signal linewidth is limited by the relatively high resistivity of refractory metals such as molybdenum or tungsten, and difficulties in forming via holes with small diameters. These factors hinder the achievement of high densities and fine patterns.

A low sintering temperature material to replace alumina in green sheets was required. A glass-ceramic material with a low dielectric constant which can be sintered at 900°C in air or in a neutral atmosphere has been found. Due to the low sintering temperature, it is possible to use not only gold conductors but also silver-palladium alloys which have far lower cost. The new multilayer glass-ceramic (MGC) substrates were developed using this low firing material.

One low firing, high dielectric constant ceramic in the $\text{Pb}(\text{Fe}_{2/3}\text{W}_{1/3})\text{O}_3-\text{Pb}(\text{Fe}_{1/2}\text{Nb}_{1/2})\text{O}_3$ binary system, has already been reported. This ceramic material can also be sintered at 900°C in air.

The new monolithic multicomponent ceramic (MMC) substrates were developed using a glass-ceramic material, a low firing high dielectric constant material, conventional metal oxide resistance materials, and metal conductors in conjunction with green sheet and cofiring technologies. Many electronic elements, such as capacitors, resistors, and wiring conductors can be incorporated in the new MMC substrates. It has been found possible to produce small, flexible, low-cost hybrid microcircuit substrates and devices.

1. Introduction

With the increasing demand for electronic equipment of reduced size and higher integration density, circuit devices

and components must have compact design and high reliability. Many electronic ceramic components, such as capacitors, resistors, inductors, substrates, and varistors can be produced with compact design and high reliability, using multilayer ceramic technology.

For logic modules in high-speed large-scale computer systems, as LSI performance and density increases, the propagation delays resulting from the length and capacitance of the interconnects on the packaging substrate can contribute significantly to the total propagation delay. To realize higher system performance, based on high-speed semiconductor LSI chips and a reduction in propagation delay time between chips, the interconnect on the packaging substrate should be designed with low resistance. For these requirements, high-density, fine patterns are necessary on signal lines at the substrate level.

In conventional packaging technology with high dielectric constant materials, the delay introduced in the package exceeds the chip delay. Therefore, it is necessary to reduce the dielectric constant for the substrate.

In the past, multilayer ceramic substrate alumina green sheets were developed which showed promise for use in high-density packaging substrates. Because this substrate sinters only at high temperatures ($>1500^{\circ}\text{C}$), it is necessary to use molybdenum or tungsten as the conductor and to sinter in a reducing atmosphere in order to prevent oxidation of these metal. In conventional packaging technology with high dielectric constant alumina ($\epsilon = 9-10$), the delay introduced in the package exceeds the chip delay. Therefore, it is necessary to reduce the dielectric constant for the substrate to below that for alumina.

Another important requirement in microelectronic packaging is small line resistance, to prevent any undesired voltage drops. To satisfy this requirement, low firing temperature multilayer glass-ceramic (MGC) substrates with gold or silver-palladium wiring have been developed. The dielectric constants for these materials are 3.9–7.8, smaller than that for alumina. These substrates thus allow high pulse transmission speed. The electrical

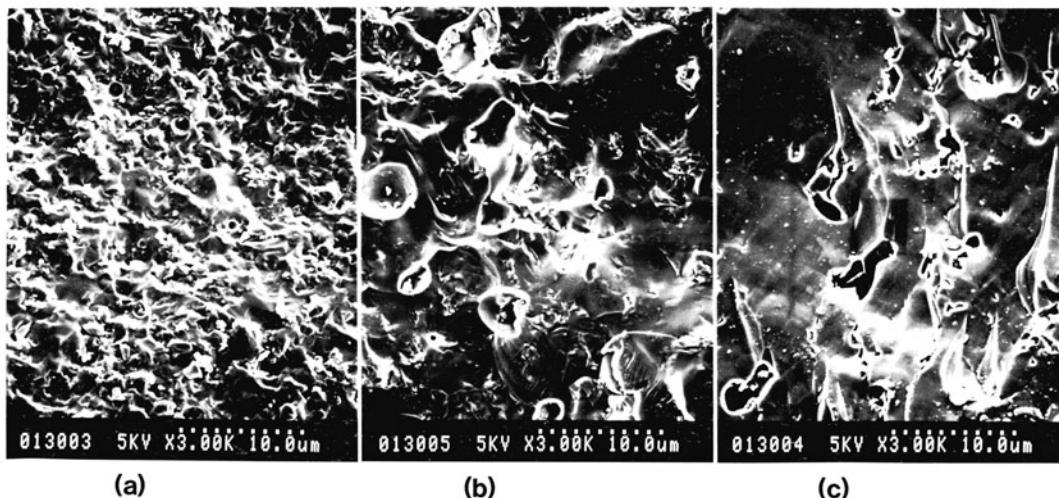


FIGURE 30 Electron micrographs of fracture surfaces of (a) Al_2O_3 -lead borosilicate glass (G-C(I)), (b) $2\text{MgO} \cdot 2\text{Al}_2\text{O}_3 \cdot 5\text{SiO}_2$ -borosilicate glass (G-C(II)), and (c) SiO_2 -borosilicate glass (G-C(III)).

resistivity of the wiring metals used in low firing glass-ceramic substrates is lower than that of molybdenum or tungsten.

Electronic circuit miniaturization is a steady trend in integrated circuit manufacturing. Hybrid microcircuits are a widely used approach to miniaturization of circuits. In a typical hybrid microcircuit, passive components are formed by thick or thin film technologies or mounted one by one as discrete components. A hybrid microcircuit can thus become large, with many components formed or mounted on the substrate. With the increasing demand for electronic equipment with reduced size and higher integration density, circuit devices and components must have compact design and high reliability.

Recently, integrated hybrid circuits, multilayer hybrid ICs with interlayered resistors and capacitors, have been developed to decrease circuit size and increase integration density.

2. Materials

a. Low dielectric constant materials. The dielectric materials used for the MGC substrates are an Al_2O_3 -lead borosilicate glass (G-C(I)), a cordierite ($2\text{MgO} \cdot 2\text{Al}_2\text{O}_3 \cdot 5\text{SiO}_2$)-borosilicate glass (G-C(II)), and a SiO_2 -borosilicate glass (G-C(III)), all having controlled particle size. The dielectric material compositions are 55 wt.% Al_2O_3 and 45 wt.% lead borosilicate glass for G-C(I), 45 wt.% cordierite ($2\text{MgO} \cdot 2\text{Al}_2\text{O}_3 \cdot 5\text{SiO}_2$) and 55 wt.% borosilicate glass for G-C(II), 35 wt.% SiO_2 , and 65 wt.% borosilicate glass for G-C(III). The optimum sintering temperature for these materials is about 900°C in air. After sintering, sample G-C(I) becomes a continu-

ous vitreous network of lead borosilicate glass, containing crystals of Al_2O_3 and devitrified glass. This glass-ceramic body has high flexural strength, ($>3000 \text{ kg/cm}^2$). Because G-C(II) and G-C(III) bodies have voids, their flexural strengths are lower ($<1500 \text{ kg/cm}^2$).

Figure 30 shows electron micrographs of fracture surface of samples of these materials, fired at about 900°C.

The frequency dependence of the dielectric constant for the three material systems are shown in Fig. 31. The dielectric constants hardly change between 1 kHz and 10 MHz, 1 MHz, values are 7.8, 5.0, and 3.9 for G-C(I), G-C(II), and G-C(III), respectively.

The dielectric constants for ceramics generally increase with increasing polarizability and atomic number, and generally decrease with decreasing bond strength and density. Highly porous structures, with high specific volume, would have low dielectric constants. G-C(II) and

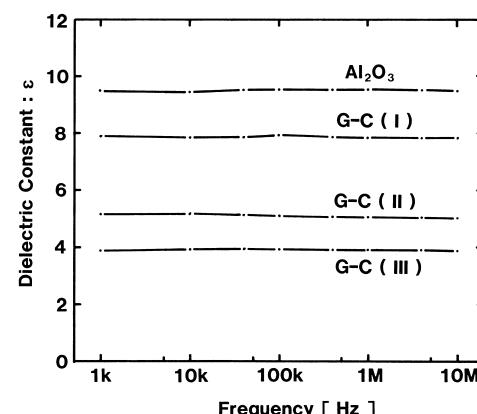


FIGURE 31 Frequency dependence of dielectric constants.

G-C(III) have extremely small dielectric constants, because of their physical properties and high porosity.

The insulator material used for the MMC substrate is GC(I).

b. High dielectric constant material. The dielectric material used for the MMC substrate is a low firing ceramic material in the binary system $\text{Pb}(\text{Fe}_{2/3}\text{W}_{1/3})\text{O}_3 - \text{Pb}(\text{Fe}_{1/2}\text{Nb}_{1/2})\text{O}_3$, with an optimum sintering temperature of around 900°C . The dielectric material composition is $\text{Pb}(\text{Fe}_{2/3}\text{W}_{1/3})_{0.33}(\text{Fe}_{1/2}\text{Nb}_{1/2})_{0.67}\text{O}_3$, the dielectric constant is high (around 7000), when processed under con-firing conditions.

c. Resistor materials. The resistor materials used for the MMC substrate are a commercial RuO_2 materials which can be sintered at around 900°C .

d. Conductor materials. Silver–palladium, silver, and gold are used for conductor wiring in the MGC substrates and MMC substrates and for internal electrodes of capacitor elements for the MMC substrate. Gold and gold–platinum are used for external terminals.

3. Structure

a. MGC substrate structure. Figure 32 shows an exploded vertical section of the MGC substrate. On the top layer, pads for assembling IC, LSI, and passive chip components are formed. The signal distribution layers are formed in the second to fourth layers. In the fifth to ninth layers, the power distribution and ground layers are formed. On the bottom layer, pads for I/O pins are formed.

b. MMC substrate structure. Figure 33 shows an exploded vertical section of an MMC substrate. On the top layer, pads for assembling IC, LSI, and passive chip

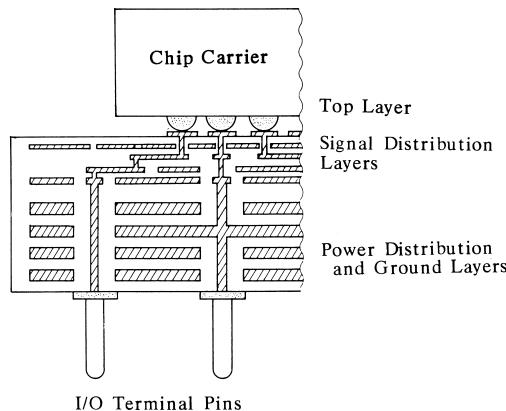


FIGURE 32 Exploded vertical section of MGC substrate.

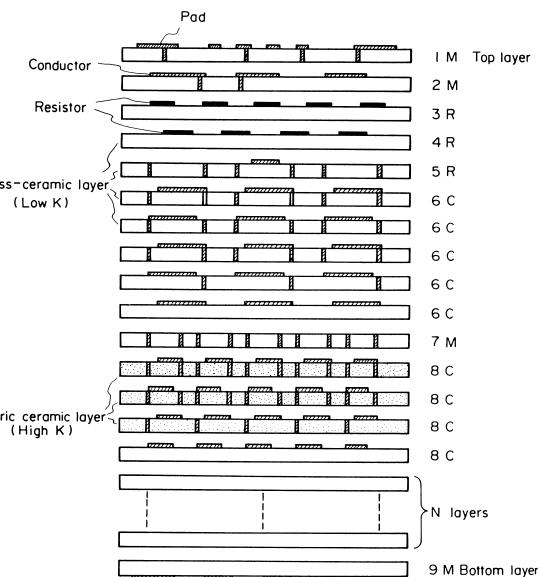


FIGURE 33 Exploded vertical section of MMC substrate.

components are formed. The second layer is the conductor layer for the signal line. Resistors are formed on the 3rd and 4th layers. Capacitors are formed on the 5th to 10th layers using the insulating glass–ceramic as the dielectric. In the 12th to 14th layers, capacitors are formed using high dielectric constant ceramic material. The layers from the 15th to the bottom, are insulator layers; they give mechanical strength to the MMC substrate.

4. Fabrication Process

a. MGC substrate process. A flowchart for the MGC substrate fabrication process is shown in Fig. 34. Lead borosilicate glass powder is ground in a ball mill until it has the desired particle size. MGC substrate characteristics, such as mechanical strength, shrinkage, and electrical properties, are strongly influenced by the glass powder particle size. Next, alumina and glass powders, with the optimal particle sizes are wet-mixed in a ball mill.

i. Slip casting. A slurry consisting of the mixed powders and a liquid vehicle (solvent, plasticizer, and binder mixture) is cast into thin ceramic green sheets ($30-200\ \mu\text{m}$) using a conventional doctor blading process. If alumina is used in fabricating green sheets, it is difficult to reduce the thickness below $200\ \mu\text{m}$. This is for convenience in handling the green sheets. Handling is improved if the percentage of organic blunder in the slurry is increased. However, an increased organic binder content deteriorates the electrical characteristics of the ceramic insulator layers, because of the pinholes which are formed

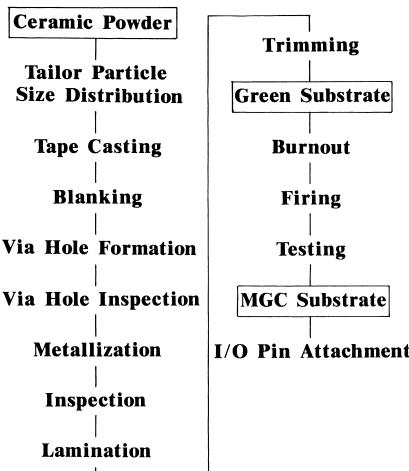


FIGURE 34 Process flowchart for MGC substrate fabrication.

during firing when only the binder is lost. In contrast, it is possible, to form thin green sheets from the glass ceramic material with no deterioration of the electrical characteristics without increasing the percentage of organic binder in the slurry.

ii. Via hole formation. The green sheets are next blanked into individual 120-mm² sheets. Via holes are formed through the green sheet by a mechanical punching process with a punch and die. It has become possible to form fine, accurately located holes through the green sheets in an economical fashion using the new punching process. The diameter of the via holes is 80 µm; a green sheet may have more than 110,000 via holes in a 100-mm² area.

iii. Metallization. Conductor paste is applied to the punched sheets using a screen printing process. The via holes are filled simultaneously with paste to form contacts between signal lines. The conductor pastes include gold, silver–palladium alloys, and gold–platinum alloys. When using gold, it is possible to print lines, owing to the low electrical resistivity of gold. Examples of printed green sheets are shown in Fig. 35.

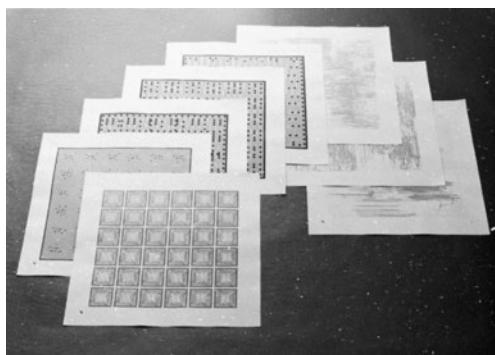


FIGURE 35 Printed green sheets for MGC substrate.

iv. Lamination. The printed green sheets are stacked in the desired sequence using the alignment holes and then laminated together at 100°C under about 250 kg/cm² pressure. Density inhomogeneities in the laminated samples influence the shrinkage of the sintered substrate. For this reason lamination should be carried out using dies and punches with flat surfaces. In addition, the stack of printed green sheets is sandwiched between two polymer sheets during the lamination to ensure uniform green density and a smooth substrate surface.

v. Firing. Sintering is performed in an electric furnace in air or in a neutral atmosphere, where conductor oxidation cannot occur. The organic binder is first volatilized at low temperatures (\approx 500°C), and then the conductor and glass-ceramic are sintered at 900°C.

b. MMC substrate process. A flow chart for the MMC substrate fabrication process is shown in Fig. 36. A slurry consisting of insulator or dielectric ceramic powder and liquid vehicles (solvent, plasticizer, and binder mixture) is cast into thin ceramic green sheets (30–200 µm) using a conventional doctor blading process. The green

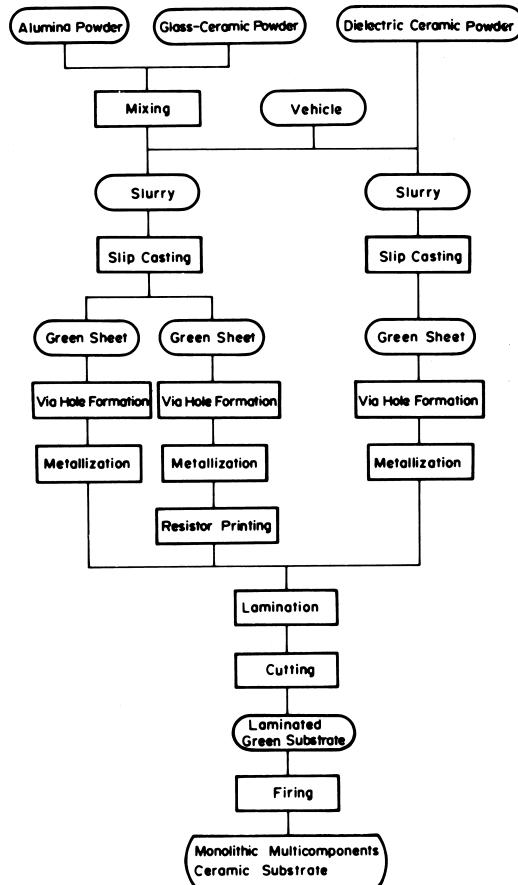


FIGURE 36 Process flow chart of MMC substrate fabrication.

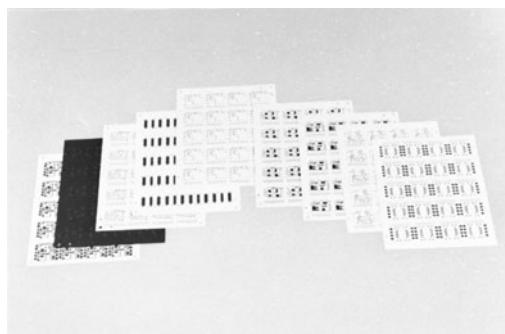


FIGURE 37 Printed green sheets for MMC substrate.

sheets are next blanked into individual 100 mm × 70 mm sheets. Via holes are formed through the green sheet by a punching process with a punch and die. In the next step, conductor paste is applied to the punched sheets using a screen printing process to form signal lines and internal electrodes for capacitor elements. Resistor pastes are applied to the metallized green sheet by a screen printing process. Examples of printed green sheets are shown in Fig. 37.

To produce the MMC substrates, the following materials properties and interactions must be carefully controlled.

1. The sintering temperature for each material
2. The shrinkage curve for each material, to achieve the same shrinkage without development of transient or permanent stresses during cofiring
3. The extent of interdiffusion at the interface between different materials during cofiring

Table VIII shows material systems for hybrid microcircuits. In the usual material system, the sintering temperatures differ for each material. Therefore, these materials cannot be readily cofired.

TABLE VIII Material Systems for Hybrid Microcircuits

Material	New system		Usual system	
	Sintering temperature (°C)		Sintering temperature (°C)	
	Material		Material	
Substrate	Glass-Ceramic	900	Al ₂ O ₃	1600
Dielectric	Low firing ceramic	900	BaTiO ₃	1400
Resistor	RuO ₂	900	RuO ₂	900
Conductor	Au, Ag-Pd	900	W, Mo	1600
			Au, Ag-Pd	900

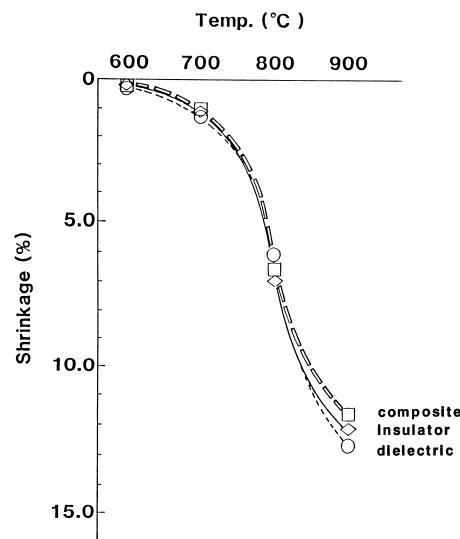


FIGURE 38 Shrinkage characteristics of the dielectric and insulator materials and of a composite substrate made from them plotted as a function of temperature.

On the other hand, in the new material system each material can be sintered at around 900°C. Consequently, a cofiring process can be used.

Figure 38 shows shrinkage curves for the dielectric material, the insulator material and a composite of these materials. These materials have almost the same shrinkage at 900°C. No cracks or delamination were found in the MMC substrates after cofiring. Accurate control of shrinkage resulted from careful control of particle size and power/green sheet processing.

Figure 39 shows a micrograph of a cross section of a MMC substrate. No significant interdiffusion can be observed at the interface between the dielectric material and insulator material. The effective dielectric constant for the dielectric material was around 7000, after cofiring.

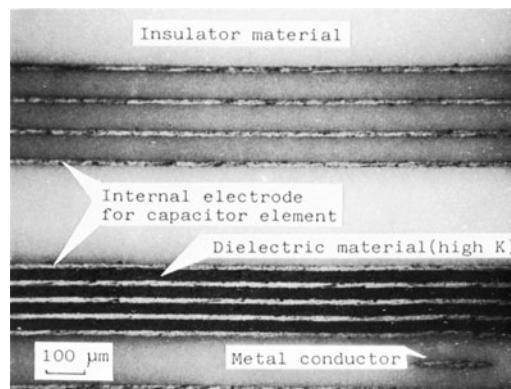


FIGURE 39 Micrograph of a cross section of MMC substrate.

TABLE IX Typical MGC Substrate Properties

Property	G-C(I) Al ₂ O ₃ -glass	G-C(II) Cordierite-glass	G-C(III) SiO ₂ -glass
Sintered density (g/cm ³)	3.10	2.40	2.15
Flexural strength (kg/cm ²)	3500	1500	1400
Camber (100 mm ⁻¹)	20 μm	40 μm	40 μm
Roughness (Ra)	0.3 μm	0.9 μm	0.5 μm
Dielectric constant (1 MHz)	7.8	5.0	3.9
Dissipation factor (1 MHz)	0.3%	0.5%	0.3%
Insulation resistance (50 V dc) (Ω cm)	>10 ¹⁴	>10 ¹³	>10 ¹³
Leak current (20 V dc) (μA)	<1	<1	<1
Thermal expansion coefficient (°C ⁻¹)	42 × 10 ⁻⁷	79 × 10 ⁻⁷	19 × 10 ⁻⁷
Linear shrinkage	13.0%	13.7%	13.9%
Shrinkage tolerance	<0.3%	<0.3%	<0.3%

5. Substrate Properties

a. *MGC substrate properties.* Typical properties of MGC substrates and glass–ceramic materials are summarized in Table IX.

i. *Physical properties.* The sintered density for G-C(I) is close to the theoretical density value, since there are very few voids, but G-C(II) and G-C(III) density values are lower than the theoretical values; because of closed pores present in these materials in the G-C(I) system, the alumina reacts with the glass phase. As this reaction proceeds, and reaction product crystallites are formed, the glass phase spreads along grain boundaries, forming a dense, continuous network holding together a dense ceramic phase. For this reason, this MGC substrate has high flexural strength (about 3500 kg/cm). Other MGC substrates, using G-C(II) and G-C(III), have less flexural strength due to the presence of voids.

Camber and roughness are extremely small on individual substrates. These MGC substrates have good flatness.

ii. *Electrical properties.* The dielectric constants at 1 MHz for the MGC substrates made from G-C(I), G-C(II), and G-C(III) ceramic powder are 7.8, 5.0, and 3.9, respectively. The MGC substrate made from the G-C(III) has much smaller dielectric constant, than conventional ceramics. The dissipation factor is less than 0.5%.

The dielectric constant and dissipation factor were measured with a multifrequency digital LCR meter, using 1 V(rms) in the frequency range from 1 kHz to 10 MHz.

The insulation resistance is more than 10¹³ Ω cm, and the interlayer leakage current is very low.

iii. *Thermal properties.* The thermal expansion coefficient is one of the most important properties of a packaging substrate. The MGC substrate for the G-C(I) system has a thermal expansion coefficient, of 42 × 10⁻⁷/°C which is constant between room temperature and 250°C, and is

close to that for the silicon semiconductor chips. The thermal expansion behavior of G-C(II) and G-C(III) remain to be optimized.

iv. *Shrinkage.* Substrate shrinkage is influenced by the particle size of the low dielectric constant material powder. The linear shrinkage values for G-C(I), G-C(II), and G-C(III) are 13.0%, 13.7%, and 13.9%, respectively. Using the technology described above, the shrinkage tolerance of cofired substrates is 0.3% or less.

v. *Substrate specifications.* Typical MGC substrate specifications are shown in Table X. The new MGC substrate can be used to produce high wiring density and fine line patterns.

b. *MMC substrate properties.* The temperature dependence of the dielectric properties of capacitors in MMC substrates based on a ceramic with the formulation is shown in Fig. 40. The dielectric constant for this

TABLE X Typical MGC Substrate Specifications

Signal lines	
Width	
Line-to-line distance	
100 μm	
200 μm	
Via holes	
Diameter	
Center-to-center distance	
100 μm	
300 μm	
Conductor resistivity	1.8–10 μΩ cm
I/O pin bonding strength	>7 kg
Green sheet thickness	100 μm
Number of layers	34

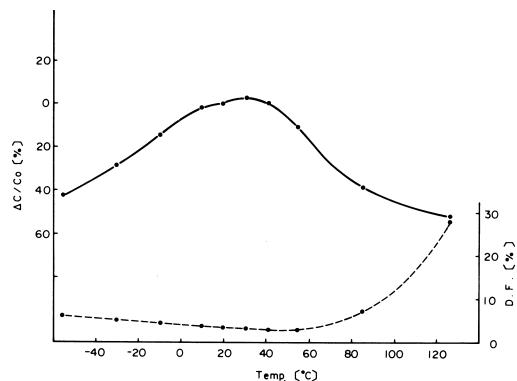


FIGURE 40 Temperature dependence of dielectric properties of capacitor elements within a $\text{Pb}(\text{Fe}_{2/3}\text{W}_{1/3})_{0.33}(\text{Fe}_{1/2}\text{Nb}_{1/2})_{0.67}\text{O}_3$ -based MMC substrate.

material is around 7000. This ferroelectric material makes it possible to form capacitors within a wide capacitance range (10 pF–3 μF) in small MMC substrates.

Typical electrical properties for the resistors in the MMC substrate are shown in Table XI. The resistivity and TCR (temperature coefficient of resistivity) depend on the conductor material and firing temperature. These resistors are fired at 900°C. Resistance elements with sheet resistivities from 100 Ω/\square to 200 k Ω/\square can be formed in MCC substrates using Dupont #1800 series pastes.

6. Applications

a. *MGC substrate application.* The production of small, high-speed, reliable computer systems requires the use of sophisticated LSI technology not only in chip fabrication, but also in packaging fabrication. The three integration scales common in large-scale computer systems are:

1. Bipolar LSI chips for low-energy current mode logic (LCML) to be mounted on multichip packages using tape-automated bonding to goldplated contacts
2. Multichip packages capable of acting as carriers for multilayer ceramic substrates or multiple LSI chips in TAB form, providing interconnection

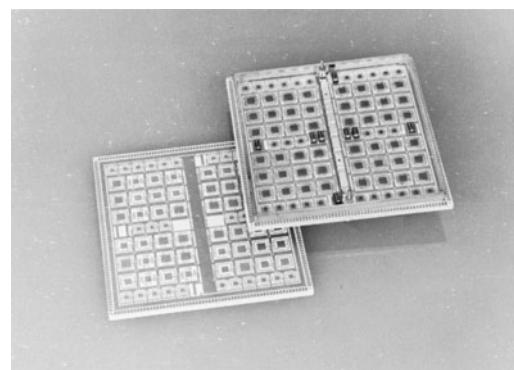


FIGURE 41 80 mm × 80 mm MGC substrates.

3. Multilayer cards for supporting and interconnecting multichip packages via multichip package connectors

Multichip packaging technology is as important to the success of large-scale processors as is LSI chip fabrication technology.

The MGC substrate is well suited as an LSI packaging substrate which can meet these requirements. The MGC substrate was experimentally applied to multilayer substrates in a multichip package for a large-scale computer system. The experimental substrate is shown in Fig. 41. The substrate has a multilayer structure produced using the green sheet method. The MGC substrate size is 80 mm × 80 mm with 240 I/O ports, 60 on each of the four edges. It houses up to 76 LSI chips on four conductor layers. TAB LSI chips are bonded to the substrate by means of an ultrasonic thermal compression bonder.

A variety of metallized green sheets are shown in Fig. 42. The patterns consist of signal wiring layers, a ground/power supply layer and chip-bonding layer. The signal line has a width of 100 μm with a minimum 2000- μm pitch. Via holes for electrical contract between layers have a minimum diameter of 100 μm . Following thick-film screen printing of gold paste and firing of samples, the sintered gold metallization resistivity turned out to be 2.4 $\mu\Omega \text{ cm}$, less than that for the same gold metallized on an alumina substrate (3.5 $\mu\Omega \text{ cm}$).

TABLE XI Electrical Properties of Resistor Elements in MMC Substrates Showing Dependence on Conductor Composition

Resistor code	Resistivity in catalog	Ag-Pd alloy (Ag/Pd = 85/15)			Au (ALP-3133)		
		Resistivity	Tolerance (%)	TCR (ppm/ $^{\circ}\text{C}$)	Resistivity	Tolerance (%)	TCR (ppm/ $^{\circ}\text{C}$)
Dupont 1821	103 Ω/\square	142 Ω/\square	8.5	$\leq \pm 130$	85 Ω/\square	7.0	$\leq \pm 210$
Dupont 1831	0.99 k Ω/\square	0.55 k Ω/\square	7.6	$\leq \pm 180$	0.24 k Ω/\square	5.8	$\leq \pm 140$
Dupont 1841	9.70 k Ω/\square	20.2 k Ω/\square	14	$\leq \pm 460$	6.4 k Ω/\square	13	$\leq \pm 140$
Dupont 1851	93.0 k Ω/\square	—	—	—	180 k Ω/\square	21	$\leq \pm 170$

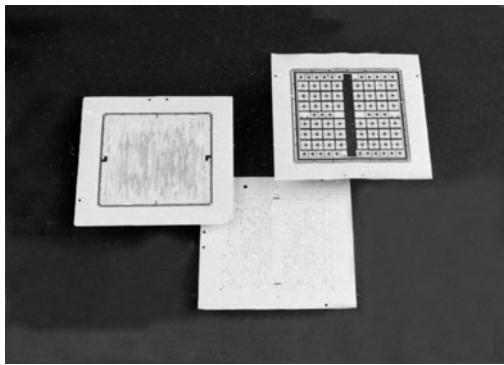


FIGURE 42 Several metallized green sheets.

b. MMC substrate applications

i. Voltage-controlled crystal oscillator (VCXO).

Voltage-controlled crystal oscillators (VCXO) are widely used as clock generators and timing signal generators in communications equipment and digital equipment. The new MMC substrate has been used as the basis for a small, inexpensive VCXO.

The VCXO circuit diagram is shown in Fig. 43. In this figure, the passive components (10 resistors and 8 capacitors) enclosed within the dotted line are incorporated in the MMC substrate.

A cross-sectional view of an obliquely polished sintered MMC substrate for the VCXO is shown in Fig. 44. Two

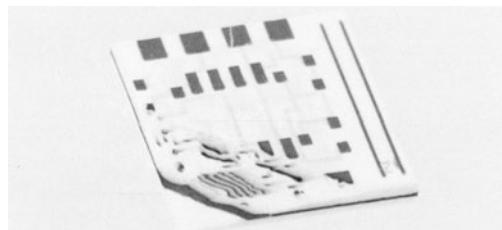


FIGURE 44 Obliquely polished MMC substrate used for VCXO.

voltage-controlled crystal oscillators, one using the MMC substrate and the other using a conventional substrate, are shown in Fig. 45. The new VCXO based on the MMC substrate oscillates at 6.4017 MHz; the dimensions of this VCXO are one-tenth that of a conventional VCXO.

ii. CR active filter. The MMC substrate has been applied to the production of small, flexible, inexpensive CR active filters.

The MMC substrate, prototype alumina substrates with conductor wiring and assembled filters are shown in Fig. 46. Ten resistors and two capacitors are incorporated in these substrates. The MMC substrates is 7 mm × 7 mm × 1 mm in size with 16 terminals, 4 on each of the four edges. For each filter, the prototype substrate has one MMC substrate, one chip containing operational amplifiers in a quad format, and five external lead frame bonding

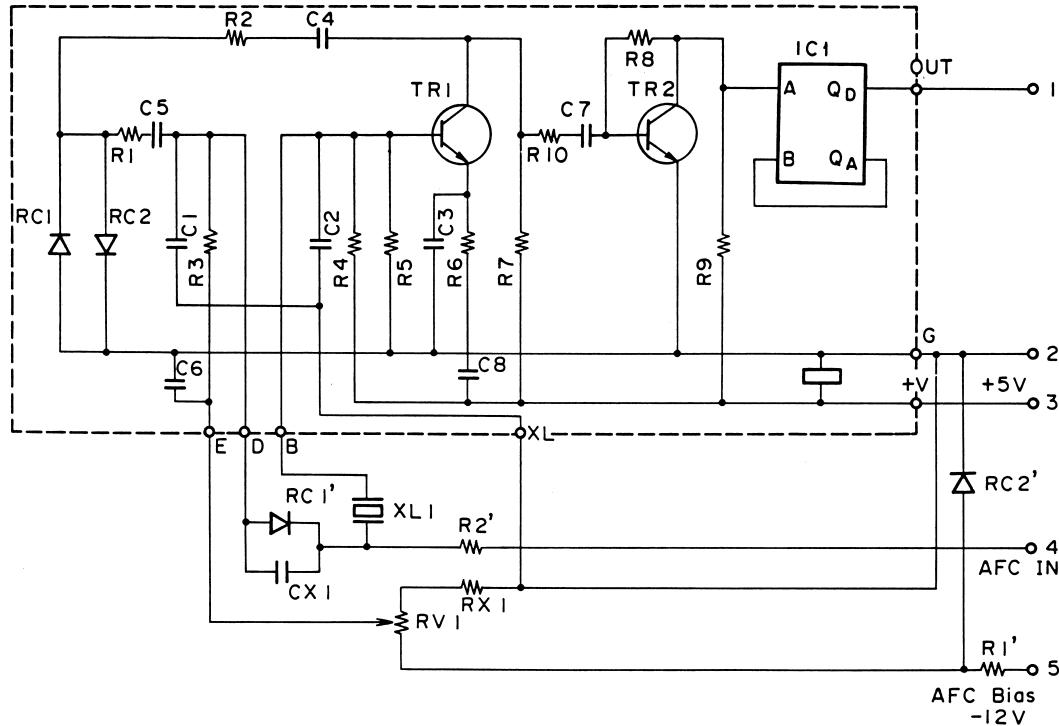


FIGURE 43 VCXO circuit diagram.

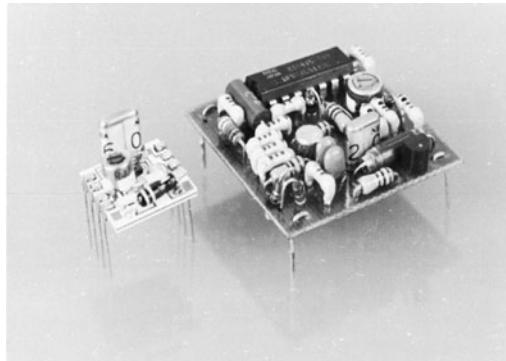


FIGURE 45 Comparison of equivalent VCXO circuits: MMC substrate and conventional substrate.

pads. Low-pass filters (LPF), band-pass filters (BPF), or high-pass filters (HPF) can be produced by mounting the same MMC substrates on different prototype substrates. These filters had good filter characteristics, which agreed well with theoretical characteristics.

iii. Memory card. The MMC substrate has been used to produce small, inexpensive 256 kbyte and 128 kbyte + 1 parity memory cards for personal computers. [Figure 47](#) shows the MMC substrate structure for the memory card. One capacitor, whose capacitance is $2 \mu\text{F}$, is formed between the ground line and the power line. [Figure 48](#) shows an assembled 256-kbyte DRAM memory card with an MMC substrate.

[Figure 49](#) shows assembled 128-bytes + 1 parity DRAM memory cards with MMC substrates in a

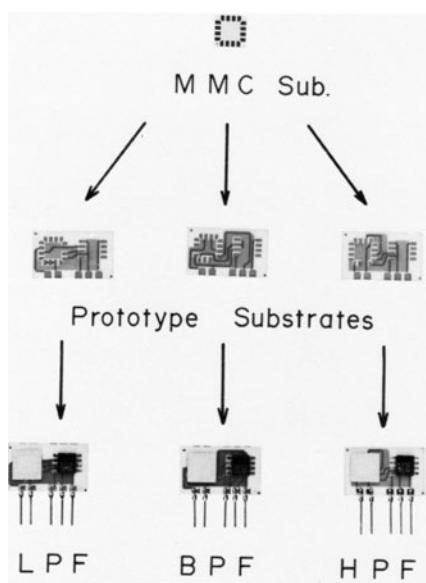


FIGURE 46 Examples of low-pass, band-pass, and high-pass filters based on MMC substrates.

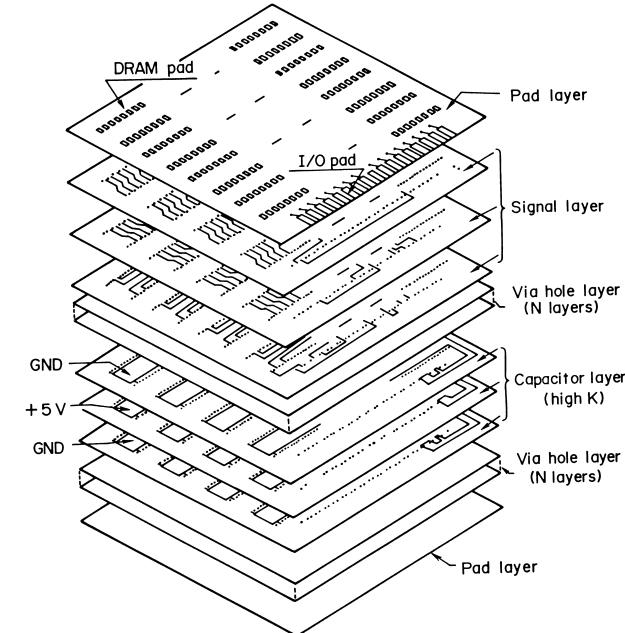


FIGURE 47 MMC substrate layer structure for the 256 kbyte memory card.

512-kbyte memory board. In each MMC substrate, one bypass capacitor for noise limiting and 24 dumping resistor elements are present.

The new memory cards are $\frac{1}{2} - \frac{1}{6}$ the size of conventional cards.

7. Future Development

Multilayer ceramic technology is progressing rapidly, especially in the field of packaging. The keys to progress in this field are advances in ceramic materials and their processing technology. Much work is needed in low firing ceramic materials to achieve lower dielectric constants and higher thermal conductivities. In addition, studies of

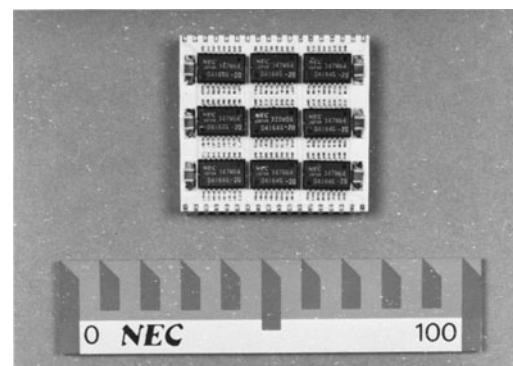


FIGURE 48 Assembled 256 kbyte DRAM memory card.

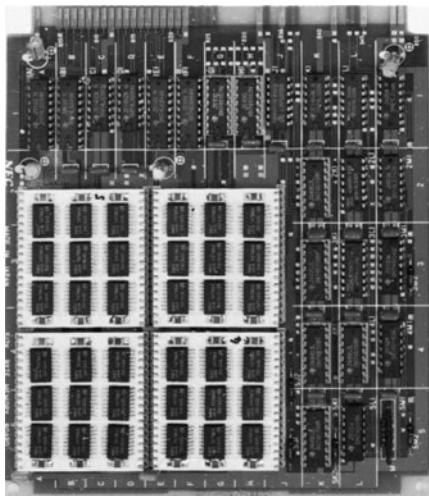


FIGURE 49 Assembled 128 kbyte memory cards on a 512 kbyte memory board.

conductor materials are required to achieve lower resistivities, processes to achieve higher wiring density for packaging substrates, fabrication of electronic components, and their assembly in hybrid components.

G. Multilayer Ceramic Substrate with Copper Wiring in Millimeter Wave Applications

1. Introduction

The development of the millimeter wave module for automotive radar, wireless LAN, and LMDS demands new substrate materials which realize low transmission loss, low conductor resistance, high reliability and low cost. For these demands, an LTCC material with low permittivity and low loss tangent has been developed for RF packages and substrates co-firable with Cu wiring. The permittivity of the material is 4.8, and the loss tangent is 8×10^{-4} at 60 GHz.

2. Material Design and Processes

To achieve low permittivity and low loss tangent, the composition of the LTCC is designed. The design concept of the new material is to minimize the lossy glass phases remained in the final crystalline phases after sintering below 1000°C.

The LTCC material consists of aluminoborosilicate glass and ceramic fillers. The glass contained SiO₂, Al₂O₃, MgO, ZnO, and B₂O₃. Final crystalline phases are spinel type, enstatite and high purity glass phase. Ingredients except for SiO₂ were removed from glass phase and used to

TABLE XII Substrate Properties

Item	LTOC	Alumina A473™	
Substrate properties			
Bulk density	g/cm ³	26	36
Thermal expansion coefficient	$\times 10^{-6}/^{\circ}\text{C}$	55	69
Young's modulus	GPa	110	270
Bending strength	MPa	200	400
Permittivity			
10 GHz r.t		48	85
60 GHz r.t		48	86
Loss tangent			
10 GHz r.t	0.0007	0.0014	
60 GHz r.t	0.0008	0.0021	
Conductor material	Ou	WMb	

construct low loss phases of spinel type and enstatite. The remaining glassy phase also turned into simple amorphous silica or pure silica glass phase of low loss tangent.

3. Materials Properties

Materials properties of the LTCC are shown in Table XII, compared with the conventional alumina (A473™). The electrical properties of the new material are superior in the characteristics of low permittivity and low loss tangent. Other properties are comparable to those of A473™ alumina.

4. Transmission Characteristics

Transmission characteristics of a microstrip line were evaluated using a HP 85109C network analyzer at various frequencies ranging from 0.045 GHz to 110 GHz. The transmission line was designed so that the characteristic impedance is equal to 50 Ω. In Fig. 50, insertion loss

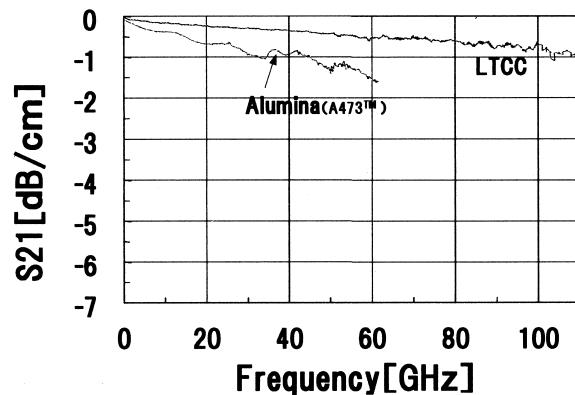


FIGURE 50 Transmission characteristic of the LTCC.

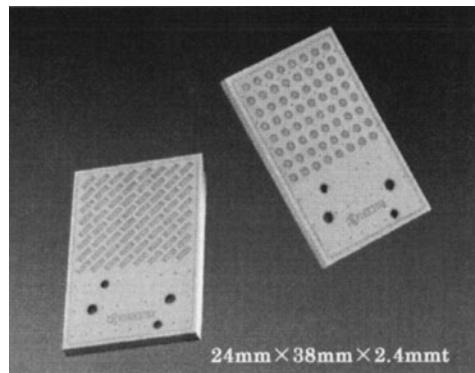


FIGURE 51 The millimeter wave antenna.

(S21) is plotted as a function of frequency. The insertion loss is 0.58 dB/cm at 77 GHz. This value is less than half of insertion loss of A473™ alumina at 77 GHz. Superior characteristics of the LTCC have been confirmed in the range of the millimeter wave frequency. The substrate system of the new material is superior in the characteristics of low permittivity and low loss tangent.

5. Application/Millimeter Wave Antenna

By combining the material and the newly developed antenna design technology, a millimeter wave antenna is now under development. The prototype of the antenna is shown in Fig. 51. The antenna characteristics were measured, the gain of this antenna was 21.5 dBi; the antenna efficiency was 42.4% at 77 GHz.

This LTCC is also used to make an evaluation sample for automotive radar and wireless LAN's millimeter wave antenna.

H. Ceramic Varistors

Hiroshi Kishi*

1. Introduction

Ceramic varistors have been widely used as electronic circuit protectors against power and voltage surges, because of their highly nonlinear current (I)-voltage (V) characteristics expressed as $I = KV^\alpha$ (α : the nonlinear exponent, K: constant). SiC varistors were developed in the 1940s and many materials such as ZnO, BaTiO₃, Fe₂O₃, and SrTiO₃ were developed for ceramic varistors after that. Table XIII shows the characteristics of these ceramic varistors. Among these varistors, ZnO and SrTiO₃ varistors having excellent nonlinear electrical characteristics based on the barrier-layer structure at the grain boundaries of the ceramics are increasingly produced.

*Taiyo Yuden Co., Ltd.

TABLE XIII Electrical Characteristics of Various Ceramic Varistors

	SiC	ZnO	BaTiO ₃	Fe ₂ O ₃	SrTiO ₃
Nonlinear exponent (α)	3 ~ 7	20 ~ 70	10 ~ 20	3 ~ 6	3 ~ 20
Withstand surge capability	High	High	Low	Low	High
Varistor voltage	5 ~ 1000	18 ~ 1800	1 ~ 3	3 ~ 30	3 ~ 1000

2. ZnO Varistor

Since the development of ZnO varistors by Matsuoka *et al.* in 1968, many studies have been done to understand conduction mechanism and to improve electrical characteristics. ZnO ceramic varistors are fabricated by sintering ZnO powders together with several kinds of additives such as Bi₂O₃, Co₃O₄, MnO₂, Sb₂O₃, and Pr₆O₁₁. The ZnO varistors consist of electrically conducting ZnO grains surrounded by electrically insulating barriers at the grain boundaries formed by the additives. In the case of the ZnO–Bi₂O₃ system, the intergranular Bi-rich phase is known as the key phase which causes the nonlinear electrical characteristics explained by the double Schottky barrier model. Their electrical properties are strongly affected by the microstructure, i.e., grain size distribution, second phases, and pores.

ZnO varistors are mainly used in circuit for protection against transient surge voltages, because of their highly nonlinear characteristics, excellent surges withstand capability and good stability.

3. SrTiO₃ Varistor

With the improvement of the surge capability of SrTiO₃-based boundary layer capacitors, SrTiO₃ varistors having capacitor characteristics were developed by Yamaoka *et al.* in 1981. The feature of SrTiO₃ varistors is their much higher apparent permittivity compared to ZnO and SiC varistors. SrTiO₃ varistors show more than 10 times the capacitance of ZnO varistors with the same size and the same varistor voltage.

SrTiO₃ varistors are fabricated as follows. Semiconductive SrTiO₃ ceramics are obtained by sintering the mixture of SrTiO₃ and donor dopants (Nb₂O₅, La₂O₃ etc.) in a reducing atmosphere. In order to make the boundary layer structure, a paste including Na₂O is printed on the surface of the semiconductive ceramics. Then the insulating barrier layer at the grain boundaries is formed by firing the printed ceramics at temperatures of 1150 to 1250°C in air.

SrTiO₃ varistors effectively absorb pulse noise, electrostatic discharge, switching surge, and indirect lightning

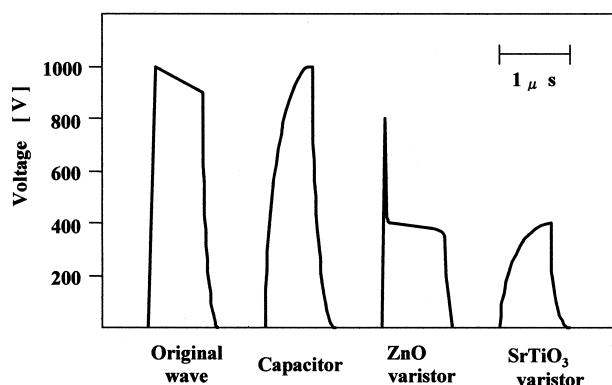


FIGURE 52 Comparison of pulse noise absorption of various devices.

stroke surge because of their large capacitance and highly nonlinear characteristics. Figure 52 shows the comparison of clamping the voltage of the capacitor, ZnO varistor and SrTiO₃ varistor, when high impulse peak voltage is applied to each device. As can be seen in Fig. 52, the SrTiO₃ varistor can effectively absorb quick rising voltage.

4. Future Trends

The technology for fabricating multilayer ceramics with thinner layers has advanced rapidly, especially in the field of capacitors. With increasing demands of down sizing and reduction of operating voltage for all electronic devices including varistors, both ZnO- and SrTiO₃-based multilayer ceramic varistors have been developed. Currently, EIA 0402 case size multilayer ceramic varistors with a varistor voltage below 10V are in mass production. Figure 53 shows the cross-sectional view of the ZnO-based multilayer ceramic varistor. To realize a lower operating voltage varistor, the control of the microstructure of the ceramic becomes more important. Further investigation is necessary in varistor materials to extend abnormal voltage coverage and withstand surge capability of smaller-size multilayer ceramic varistors.

Bibliography for Section I.H can be found at the end of this article.

II. STRUCTURAL CERAMICS

A. Heat-Resistive Ceramics for Engines

Makoto Yoshida*

Heat-resistive ceramics have been developed for energy conservation and environmental protection. The broad scientific and engineering activities in many countries have led to new materials with superior mechanical properties

*Kyocera Corporation.

and improved environmental and practical performance. The development of high-strength materials used under combustion condition of engines was conducted for increasing fuel efficiency and output power, quick starting, decreasing the weight, prevention of noise, increasing abrasive resistance, etc. Now there is a demand for the engines used for automotive vehicles and power generation systems, which have higher efficiency, better protection of the environment, and lower weight.

1. Ceramics in Reciprocating Internal Combustion Engine

Much of the early work on the application of ceramics to reciprocating engines concerned cylinder insulation and diversion of heat from the coolant to an energy recovering devices in the exhaust pipe, as a means of increasing overall fuel efficiency. In these early days, however, ceramic components were unable to withstand such an application, since the level of development of ceramic material group was not high enough. Ceramic portliner is the only structural component which cleared practical application for the heat insulation since 1986. The part increases the temperature of the exhaust gas to get higher efficiency of turbocharger and catalytic converter. The portliner is made of aluminum titanate and has a high thermal shock resistance.

In the early 1980s, many ceramic components were developed for the heat engine application as the result of material development for heat insulation engines. The material properties required for the engine components are

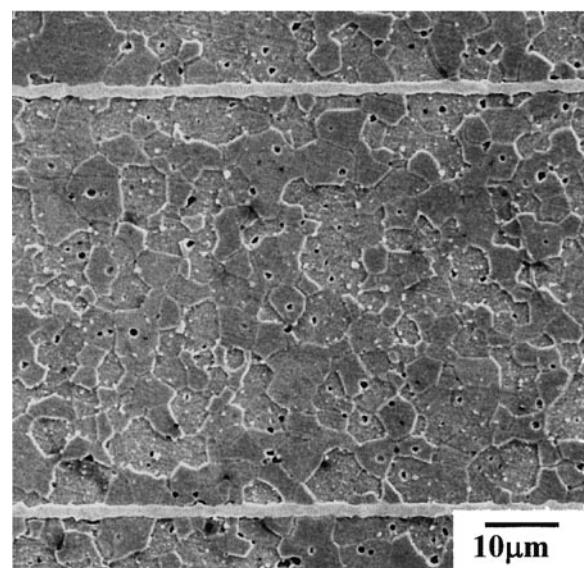


FIGURE 53 SEM photograph of the cross section of ZnO-based multilayer ceramic varistor.

- (1) Rocker arm Tip (Valve side)*
- (2) Adjusting Seat
- (3) Valve Guide
- (4) Exhaust Portliner*
- (5) Valve
- (6) Turbocharger Rotor*
- (7) Valve Seat
- (8) Cylinder Liner
- (9) Piston Crown
- (10) Piston Pin
- (11) Rocker arm Tip (Push Rod side)*
- (12) Push Rod Tip
- (13) Swirl Chamber (Upper)*
- (14) Glow Plug*
- (15) Swirl Chamber (Lower)
- (16) Cylinder Head
- (17) Tappet*
- (18) Cam

*Under Mass Production

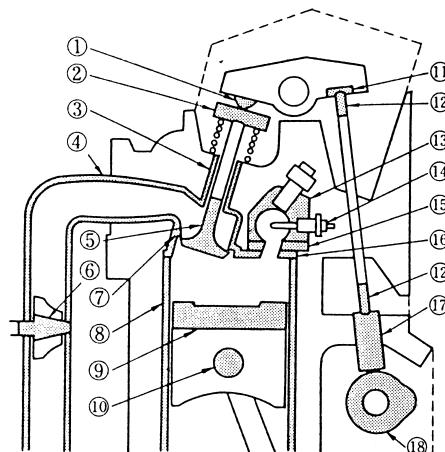


FIGURE 54 Ceramic components under development for the reciprocating engine.

of high strength at elevated temperatures, thermal shock resistance, fracture toughness, corrosion resistance, wear resistance, and light weight. Among many engineering ceramics, silicon nitride, silicon carbide, and partially stabilized zirconia were intensively investigated to clear these requirements. Silicon nitride had been mainly selected for such ceramic engine components under mass production. Figure 54 shows the possible application of ceramic components for reciprocating internal combustion engines, and an asterisk indicates the ceramic components in practical use.

The first commercial product for engines made of silicon nitride ceramics was the glow plug, and it has been manufactured since 1981. The ceramic glow plug is suitable for quick starting of the diesel combustion engine. Silicon nitride was selected for its high thermal shock resistance. A hot pressing process was used for the densification of silicon nitride with embedded metal tungsten filament.

The swirl chamber is used for indirect injection of the diesel engine. A particularly attractive goal is reduction of particulate emissions and increasing output power due to higher combustion efficiency expected at higher operating temperatures. The swirl chamber made of conventional nickel-based superalloy was oxidized and sometimes cracked under such temperature conditions. Silicon

nitride, instead of the heat-resistive alloy, was selected for the superior thermal fatigue resistance.

In 1986, a silicon nitride turbocharger rotor was developed for reducing the turbo lag and increasing the acceleration response. Since the ceramic rotor has only 40% the weight of the metallic one, the inertia moment of the ceramic rotor decreases to 60% of the metallic one. The bending strength of silicon nitride ceramics for the turbocharger rotor is nearly 1000 MPa at room temperature. There was much improvement in the reliability of ceramic material during the development of ceramic rotors.

Rocker arm tips, tappets, and cam roller followers were developed for wear resistance of ceramic material. Recently, the high-pressure fuel injection system has been developed for the exhaust emission control, especially for particulate material reduction. Silicon nitride is used as wear-resistive material for the high-pressure fuel pump. Under these conditions, the wear volume of silicon nitride components was decreased to 60% less than that of the metallic ones.

2. Ceramics in Gas Turbine Engine

Gas turbines have many advantages such as light weight, small size, multifuel capability, low emissions, low noise, and low vibration. However, the shortcoming is that their

thermal efficiency decreases as their output power decreases. The thermal efficiency of small-size engines under 500 kW is 15~20%. In order to improve the thermal efficiency, the turbine inlet temperature should be increased. There are, however, some difficulties for increasing the TIT when metal components are used. It is also difficult to apply internal cooling systems for small metal components because of their size limitation. For this reason, ceramics have been actively investigated as heat-resistant components in the gas turbine engines for high thermal efficiency. In recent years, it has become very important to ensure a stable supply and effective utilization of energy resources, and to find a solution for environmental pollution problems. In this direction, the ceramic gas turbine is also believed to be one of the most promising candidates.

During the 1960s and 1970s, the silicon-based ceramics were further developed for the gas turbine engine application. For silicon carbide, the major consolidation processes are reaction bonding, hot pressing, and sintering. The use of boron and carbon as sintering additives was a particularly noteworthy advance. For silicon nitride, the major consolidation processes were reaction bonding and hot pressing. Special progress was made with the addition of various oxides such as magnesia, alumina and yttria which were necessary for densification by the hot pressing process.

During the 1980s and 1990s, the sintering process of silicon nitride was further developed for complex shaped components. Gas pressure and normal pressure sintering processes were adopted for the fabrication of gas turbine components. **Figure 55** shows the typical silicon nitride components for a 300-kW class gas turbine engine. Silicon nitrides are usually sintered with an addition of metal oxides, especially with rare earth oxides, in order to reduce sintering temperature. There are many different types of crystalline phases at grain boundaries. Mechanical and chemical characteristics of sintered silicon nitride materi-



FIGURE 55 Ceramic components for 300-kW class gas turbine engine.

als depend greatly on the stability of these grain boundary crystalline phases. For example, a silicon nitride with an oxide-type grain boundary, such as rare earth disilicate, has higher oxidation resistance than that with an oxinitride-type grain boundary, such as YAM and wolastonite. As for high-temperature strength of the sintered silicon nitride, there is a strong relationship with the melting point temperature of the grain boundary phase. It was observed that both the flexural strength at 1400°C and the melting point temperature of the rare earth disilicate increase as the rare earth ionic radius decreases. These results clearly indicate the way for obtaining the high heat-resistant material, suitable for the ceramic gas turbine application. **Figure 56** shows the improvement of the flexural strength of silicon nitride ceramics for the past 20 years.

Structural ceramic components for engines are still in development, and material properties are making steady progress. They are the key components for resolving the global problem, such as fuel consumption, emission, and noise.

B. Ferrite Materials

Taku Murase*

Naoyoshi Sato*

Atsuyuki Nakano*

Hitoishi Taguchi*

Takeshi Nomura*

Recently, with the progress and development of mobile devices, demands for ferrite components are small sizing and high efficiency. The key to obtain high-performance soft and hard ferrites is to investigate the means of high-technology processing. The current research is reviewed with low core-loss MnZn-ferrite, multilayer chip ferrite, and M-type Sr-ferrite magnets.

1. Introduction

Ferrites are one of the most essential materials in the electronic industry. It is generally possible to classify them into soft and hard ferrites. In the case of soft ferrite materials, the coercive force is small, and saturation magnetization is filled with a comparatively small magnetic field. On the other hand, hard ferrites have high coercive force and residual flux density, and maintain the spontaneous magnetization. In other words, soft ferrites are attracted to permanent magnets, and hard ferrites are permanent magnets. Representative soft ferrites are MnZn-ferrites and NiZn-ferrites, and typical hard ferrites are Sr-ferrites and Ba-ferrites. Soft ferrite materials are produced for transformers and inductors at higher frequencies. Hard ferrite materials have been used extensively for acoustic devices and motor magnets.

*TDK Corporation.

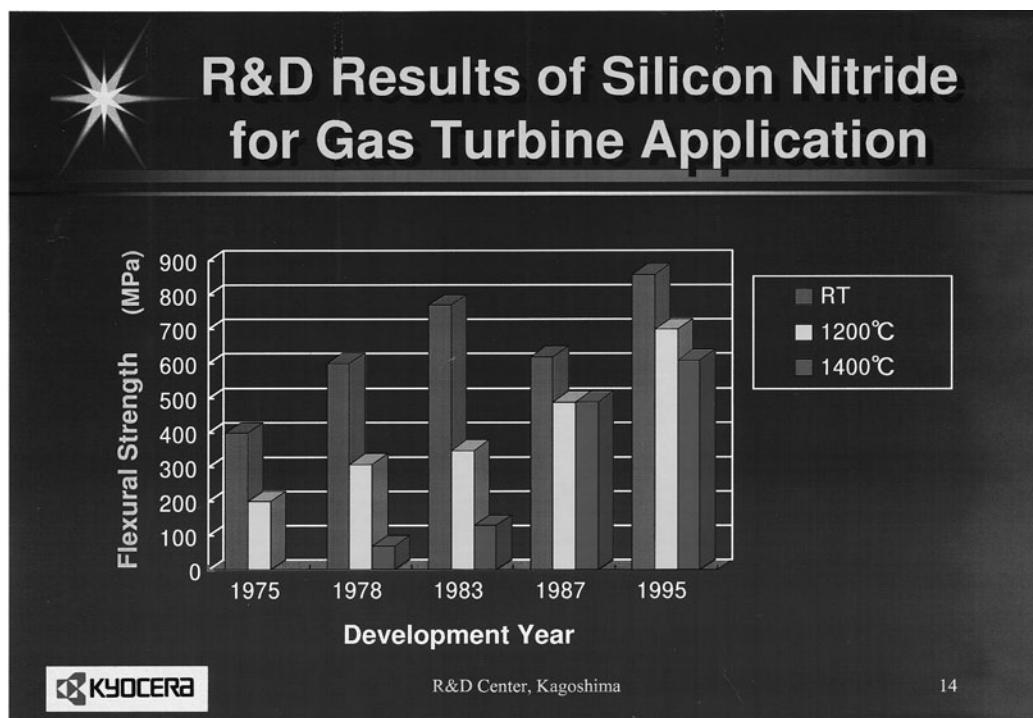


FIGURE 56 Improvement of the flexural strength of silicon nitride ceramics for the past 20 years.

MnZn-ferrites comparatively have the highest saturation magnetization and permeability in other spinel ferrites. The MnZn-ferrite materials with higher permeability in a small magnetic field are mainly being used for filters, wide band, or signal transformers. Materials with higher saturation magnetization and lower core loss in a comparatively strong magnetic field are being used for power supply and choke coil cores. The current research has concentrated on the low core-loss MnZn-ferrite for power applications (Nomura *et al.*, 1982; Nomura, 1992; Ochiai and Okutani, 1985; Otsuki, 1992). A higher operating frequency of small-sized switched-mode power supplies is imperative to enhance their efficiency. When the driving frequency is increased from hundreds of kilohertz to several megahertz, the core loss of MnZn-ferrites is lower than that of other ferrites and alloys. The properties of NiCuZn-ferrites are designed for applications in the radio wave band. The good point is that the cores can be directly wound with wires, because the materials with poor Fe_2O_3 have high resistivity. NiCuZn-ferrite materials are used for rotary transformers, inductors, and power materials for signal processing within various electronic equipment. Recently, NiCuZn-ferrites are necessary materials for EMC components. The EMC components are useful on mobile communication and high-speed personal computers. Since the unnecessary electromagnetic noise has caused large problems on precision apparatus and as-

sociated devices, the regulations for the acceptable noise level in some countries has been enforced to higher frequencies. In order to solve this noise problem, it is effective to use magnetic loss which occurs at the resonance frequency of polycrystal ferrites. The EMC components have been produced for multilayer chip inductors (Nomura and Takaya, 1987; Takaya, 1988) as well as for small cores. The multilayer ferrite chip inductors are indispensable for miniaturizing devices. Hard ferrites are produced in large quantities for motor magnets. One of the typical applications is for automobile parts such as scooters, wipers, and fans. The trend has been toward miniaturizing and weight reduction. Requisite properties for hard ferrite materials are higher efficiency and reliability. Consequently, high-performance Sr-ferrite magnets (Cochardt, 1963) are mainly produced. One of the processing techniques is used to wet-press in a magnetic field (Fisher, 1983; Stuijts *et al.*, 1954). A new composition, lanthanum and cobalt substitution in M-type hexaferrite, was developed with extremely good properties (Iida *et al.*, 1999; Nishio *et al.*, 1999).

The production of ferrites is used for “ceramic method,” in which the powders are shaped and sintered. The electromagnetic properties are influenced by the microstructure as well as by the chemical composition. In order to achieve high performance, it is necessary to introduce the excellent processing techniques for obtaining uniform microstructures and to improve the just analysis for characteristics.

TABLE XIV Magnetic Properties of MnZn Ferrite Materials

Characteristics	PC40	PC44	PC45	PC46	PC50
μ_i	2300	2400	2500	3200	1400
Core loss [kW/m ³]	600 (25°C)	600 (25°C)	570 (25°C)	350 (25°C)	*130 (25°C)
at 100 kHz, 100 mT	450 (60°C)	400 (60°C)	250 (75°C)	250 (45°C)	*80 (60°C)
* at 500 kHz, 50 mT	410 (100°C)	300 (100°C)	460 (100°C)	660 (100°C)	*80 (100°C)
Bs [mT]	510	510	530	530	470
Br [mT]	95	110	120	80	140
Hc [A/m]	14	13	12	11	37
Tc [°C]	>215	>215	>230	>230	>240

The aim of the present paper is to focus on the control of magnetic properties and microstructures in ferrite materials.

2. MnZn-Ferrites

MnZn-ferrites have a comparatively high-saturation magnetization. Furthermore, there is the specific chemical composition such that the values of magnetic anisotropy and magnetostriction are reduced to zero (Ohta, 1963). It is possible to form fast-moving domain walls in order to control the microstructure (Roess *et al.*, 1961, 1964). Consequently, the MnZn-ferrite permeability is the highest in other ferrites with spinel structure.

a. Production process for MnZn-ferrites. The manufacturing process for MnZn-ferrite cores requires the sintering process to be carefully controlled with the atmosphere. The difference between MnZn- and NiZn-ferrites is apparent in the manufacturing process. When the mixed oxide of the MnO–ZnO–Fe₂O₃ system is sintered in high-oxygen partial pressure, manganese oxide does not result in divalent oxide (MnO), and forms partially the trivalent oxide (Mn₂O₃) within the ferrite. Furthermore, in the case of an Fe₂O₃-excess composition, iron oxide creates the divalent ions (Fe²⁺) at high sintering temperatures, and changes into maghemite (γ -Fe₂O₃) which has cation vacancies within the ferrite, or deposits hematite (α -Fe₂O₃) by cooling in high-oxygen partial pressure (Blank, 1961). Furthermore, the suitable additives of SiO₂ and CaO are effective in MnZn-ferrite applications in order to lower the loss (Akashi, 1961), because these additives form the grain-boundary layers of high resistivity. The effectiveness of several dopants other than SiO₂ and CaO is well known. For example, TiO₂ or SnO₂ additives are used to lower the core loss.

b. Power ferrites. The switched mode power supplies are equipped in almost all of the electronics devices. The production of MnZn-ferrite materials for power

applications is increasing with the progress of mobile devices. The strong demands of electronic components are miniaturizing and higher efficiency. Consequently, MnZn-ferrites for power applications are required to reduce core loss. In recent years, MnZn-ferrite materials have achieved lower losses by selecting compositions and additives, and controlling the grain size and the grain boundaries. Table XIV shows the electromagnetic properties of MnZn-ferrite for power supply and choke coil cores. The PC40 has been produced since 1982, and the others have been developed since 1992. The core loss at 100 kHz has decreased to 110 kW/m³ for 10 years. The successful approaches were to use high-purity raw materials, which have lower phosphate, and to avoid the defects with the sintering atmosphere.

It is well known that additives of SiO₂ and CaO are important to obtain low loss, because the grain-boundary layer forms with the additives and influences the electromagnetic characteristics. In recent years, the current research has concentrated on the relation between grain-boundary chemistry and sintering conditions (Franken and Stacy, 1980; Paulus, 1966; Sato and Otsuki, 1996). The sintering schedules and microstructures of MnZn-ferrites are shown in Figs. 57 and 58, respectively. Microstructures were observed by optical microscopy after polishing and etching in HF. In the case of N₂-annealing, the grain boundary is hardly clear. It is suggested that the elements of Ca contaminate the ferrite-grains. When the forms of grain boundary are imperfect, the permeability and resistivity lower remarkably, and the core loss increases. The key to achieving lower loss is to control grain-boundary chemistry as well as the divalent iron-ions and cation vacancies.

The usual way to reduce the size of switching power supplies is to increase the operating frequency. The PC50 is material designed for power applications at higher frequencies over 500 kHz, as shown in Table XIV. In order to improve the core loss at higher frequencies, the composition and additives are different from the material for 100-kHz applications. The ferrite developed for

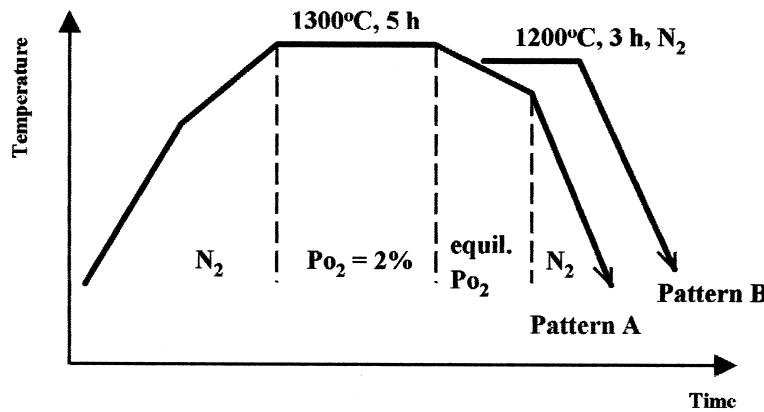


FIGURE 57 Sintering schedule.

high-frequency applications has a uniform microstructure with a grain size of about $5\ \mu m$. It is supposed that the fine microstructure reduces the range of an eddy current and leads the resonance frequency to increase. The results suppress eddy current and residual losses with increasing frequency. The key for high-frequency applications is to produce high-density as well as small grains.

3. NiCuZn-Ferrites

The incessant demand for higher-density circuits in electronics during recent decades has required the continued miniaturization of all components. Surface mount technology (SMT) was developed resulting in densification of component mounting. Improvement of SMT and SMD (surface mount devices) gave revolution of the electronics. On the other hand, “noise” at radio frequencies is generated by the high-density circuits and microprocessor-controlled devices in the electronics. Today ferrite products especially multilayer chip ferrites are a common component used to attenuate Electromagnetic Interference (EMI).

Multilayer chip ferrite components were developed using thick-film printing and co-firing technology (Nomura and Takaya, 1987; Takaya, 1988). Figure 59 illustrates a structural model of the multilayer chip inductor. Silver is used as the internal winding material and low-temperature sintered NiCuZn-ferrite is employed as an inductor material due to its co-firability with Ag. The multilayer chip inductor will be a milestone in the history of ferrites.

a. Features and processing of the multilayer chip ferrites. In the multilayer chip inductors, magnetic flux goes through the ferrite, so the magnetic flux does not leak out from the chip. Moreover, the ferrite forms a magnetic shield against external magnetic fields. Therefore, the multilayer chip inductor is especially suited for higher-density circuits in electronics. The main features of these composite multilayer chip components are as follows. (1) Sizes are easy to handle by the mounting machine. (2) They are highly dense and versatile. (3) They have high reliability due to a complete monolithic construction. (4) There is

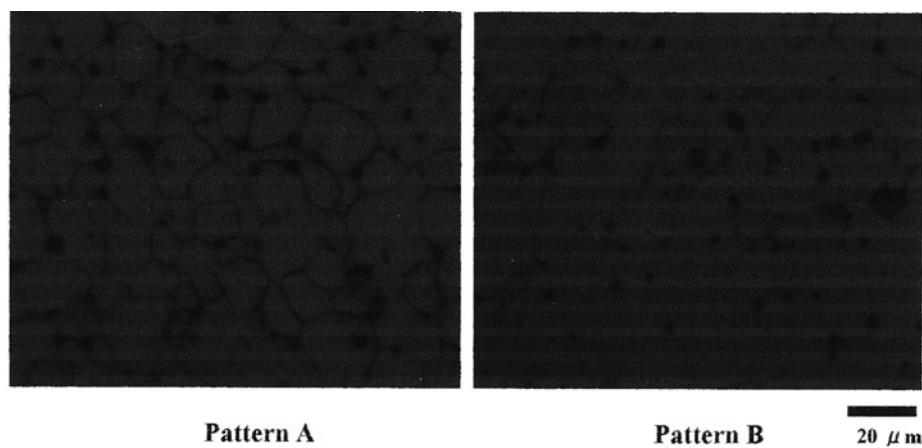


FIGURE 58 Microstructures of MnZn ferrites.

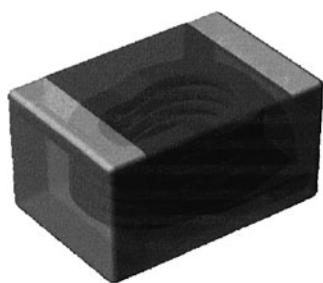


FIGURE 59 Structural model of the multilayer chip inductor.

no cross talk due to magnetic shielding. (5) High-density mounting is available. With regard to the size of the multilayer chip inductors, the chips are available in a wide variety of sizes, including the miniature-type 1005 (length: 1.0 mm, width: 0.5 mm). The multilayer chip ferrite components are made using screen printing or sheet stack process. The printing process of multilayer chip inductors is shown in Fig. 60. First, ferrite paste is printed up to about $300\ \mu$, and then silver paste is printed to form a half-turn of winding. The ferrite paste is next printed to hide the silver layer, except for the connection part of winding. The winding is made up alternately of the ferrite paste and the silver paste in this manner. The thickness of the ferrite layers is designed to be about 15–40 μ . (Kanagawa *et al.*, 1988; Nakano *et al.*, 1990).

Usually ferrite materials are sintered at about 1200–1400°C, which is quite high. In order to obtain a practical *Q* factor of multilayer chip ferrite, the winding material should be 100% silver internal electrode, therefore low-temperature sintering of ferrite is necessary. The requirements for ferrite materials of multilayer chip ferrite are

as follows (Smit and Wijn, 1959): (1) sintering temperature less <950°C due to internal Ag conductor, (2) no reactivity with internal Ag conductor, (3) high resistance due to co-firing with internal Ag conductor, (4) high *T_c* to withstand solder temperatures without adverse effects, and (5) high initial magnetic permeability. It is known that NiZn-ferrite has high resistance, and that CuZn-ferrite is a low-fire composition, therefore NiCuZn-ferrite was studied to combine both properties. The higher the cupric oxides content the lower the sintering temperature. This is due to the fact that the magnetic properties of NiCuZn-ferrite depend strongly on composition, and the final composition can not be selected from this viewpoint alone. Consequently, cupric oxide content should be as high as possible without compromising the magnetic properties. The sizing of ferrite particle is also a very important technique for increasing densification. In accordance with the sintering theory, the densification temperature of NiCuZn-ferrite decreases with increasing the specific surface area of ferrite powder. When the specific surface area of the NiCuZn-ferrite powder is over $7\ m^2/g$, the ferrite powder can be co-fired with silver under 900°C. To achieve over $7\ m^2/g$, techniques were developed to adequately extend the milling time while minimizing contamination from the media (Smit and Wijn, 1959).

b. Co-firing with Ag and ferrites

i. *Stresses from Ag internal electrode.* In order to achieve the multilayer chip ferrites many studies for co-firing technology of ferrite and Ag were required. In a previous study it was reported that the Ag conductor gave rise to significant stresses which had a deleterious effect on the magnetic characteristics of sintered chips. It is well

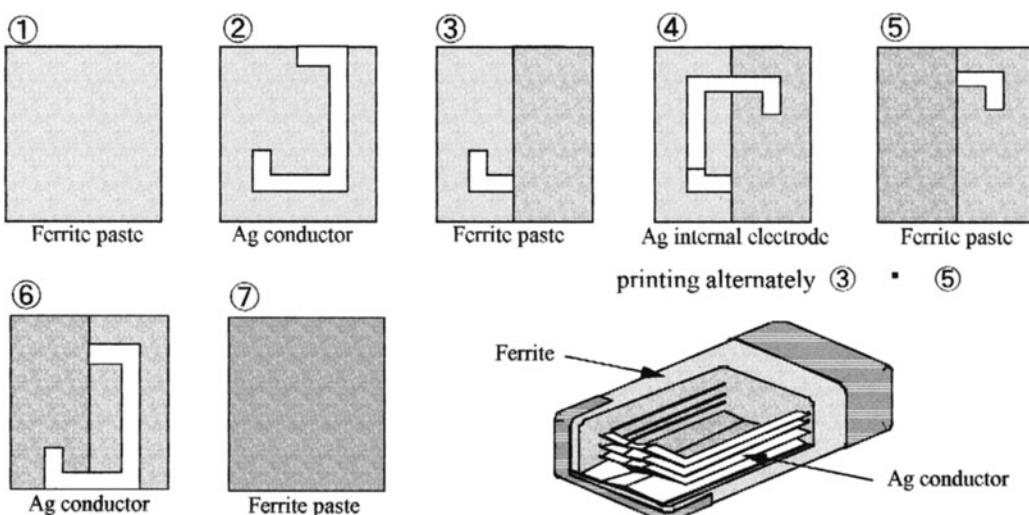


FIGURE 60 Printing process of multilayer chip inductor.

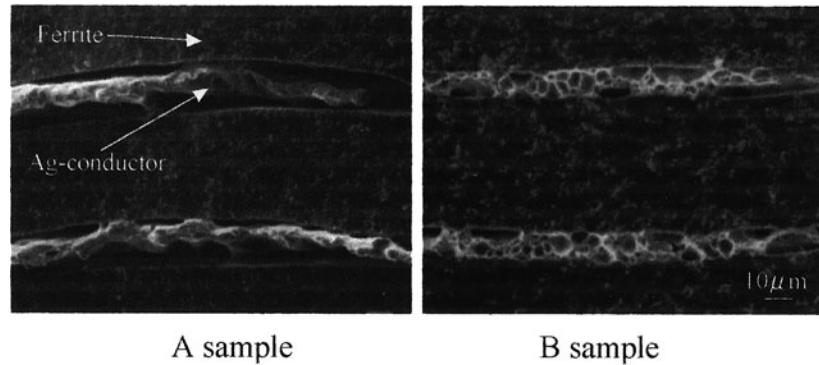


FIGURE 61 Typical SEM's of fracture surface of chip inductors.

known that the magnetic characteristics of ferrites depend on their composition, additives, microstructure, and residual stresses (Takaya and Nomura, 1994). Particularly, the permeability of the magnetic characteristics is strongly sensitive to stresses. In order to obtain high performance of multilayer chip ferrites, it is very important to control residual stresses between the ferrite materials and the Ag. The thermal expansion coefficient of Ag is about two times that of NiCuZn-ferrite. Differences between the thermal expansion coefficients of conductor material and ferrite material should be as small as possible in order to restrain the deterioration of magnetic properties. Figure 61 shows SEM micrograph of a typical chip inductor fracture surface. The electromagnetic properties of multilayer chip inductors containing different shrinkage of Ag conductors were compared. The large-shrinkage Ag conductor with lower contact of Ag conductor to the ferrite layer as seen in A resulted in better magnetic characteristics than the chip inductor using a small-shrinkage Ag conductor. It is thought that the lower contact of the Ag conductor to the ferrite layer can reduce the overall residual stress from the Ag conductor. Adequate control of the condition of the Ag conductor will be key to achieving high performance in multilayer chip inductors (Nomura and Nakano, 1992; Takaya and Nomura (1994)). Recently, a new approach of using low magnetostriction in ferrites has been investigated in order to achieve high chip performance. According to the equation for permeability, even if stresses were high, permeability would be higher if magnetostriction were kept low. It is known that magnetostriction of Mg-ferrites is lower than that of Ni-ferrites. When each ferrite composition is the same, the magnetostriction of MgCuZn-ferrites is about one fourth of the magnetostriction of NiCuZn-ferrites. It was expected that the ferrite materials have high potential in multilayer chip ferrite component applications, not only in chip inductors but also in LC and EMI chip filter components. The effect of compressive stress on change in μ of MgCuZn-ferrite is shown in Fig. 62. It is known that compressive stresses

cause a decrease in μ of both materials. However, the change in μ of MgCuZn-ferrite is smaller than that in the NiCuZn-ferrite. Figure 63 shows the effect of co-fired with dielectric material on change in μ of MgCuZn ferrite. In this experiment, when the green body was pressed, dielectric material was also pressed with the ferrite. Then the samples were fired. The dielectric material is the low-temperature sintering material which is TiO_2 added CuO 5wt%. It was expected that stress occurred from the interface between dielectric material and ferrite material. In the case of the co-fired sample of NiCuZn-ferrite, the μ shows a -22.9% decrease. On the other hand, μ of the MgCuZn ferrite sample decreased only -3%. These results are consistent with our assumption. Therefore, it is thought that MgCuZn-ferrite materials have a high potential for multilayer chip ferrite components (Nakano and Nakah, 2000).

ii. Ag diffusion. Another deleterious effect is thought to be the diffusion of Ag into the ferrites, causing a decrease in the inductance of the chips. A study of the effect of Ag-diffusion into ferrite in order to control it is essential for successful co-firing. When the mechanism of

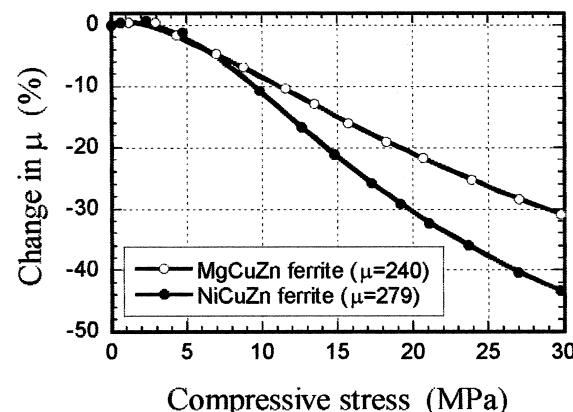


FIGURE 62 Effect of compressive stress on change in μ of MgCuZn ferrite.

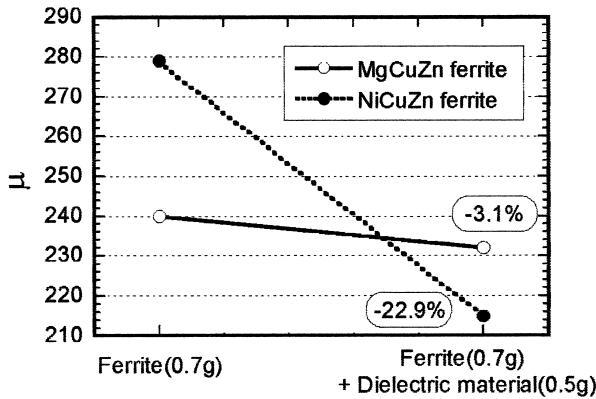


FIGURE 63 Effect of cofired with dielectric material on change in μ of MgCuZn ferrite.

Ag-diffusion was investigated it showed that sulfur and chlorine ions react with Ag during heating, then combine with Ag to give Ag_2S and AgCl . They are very low melting temperature materials (about 450–600°C), which therefore promote Ag diffusion into a ferrite body (Smit and Wijn, 1959). The effect of sulfur content on the Ag-diffusion behavior in the multilayer chip inductor was investigated and is shown in Fig. 64. It is observed that there is less Ag in ferrite containing 170 ppm of sulfur than in the sample with 680 ppm. The effect of Ag on the permeability of NiCuZn ferrite was studied. The permeability decreased with increasing Ag. The Curie temperature of these samples was also measured, and was found to increase with increasing Ag. NiCuZn-ferrites generally contain sulfur and chlorine, which are impurities of Fe_2O_3 and NiO . Ag-diffusion can be controlled by the selection of raw materials for ferrites (Nakano and Nakah, 2000).

iii. Stresses from grain boundary. TEM investigated sintered NiCuZn-ferrites of multilayer chip inductors at 890°C (fired). It was observed that a great number of interference fringes existed at the grain boundary. It seems that the impurities existing at the grain boundary brought

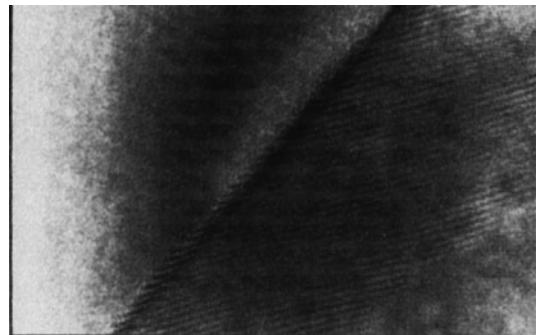
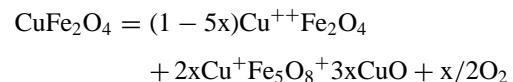


FIGURE 65 A lattice fringe image of the interference fringes.

about significant stresses, which caused the interference fringes. A lattice fringe image of the interference fringes has been studied by TEM as shown in Fig. 65. At the interference fringes, a low-angle tilt boundary was formed and the d-spacing narrowed (Nakano *et al.*, 1992). From this result, it is certainly known that significant compressive stress exists in the ferrite grains. A TEM-EDX investigation was performed to identify the precipitation at the core of the interference fringes and at the triple point of the ferrite body, showing a high concentration of CuO_{1-x} and Ag (Nakano *et al.*, 1992). It is well known that CuFe_2O_4 dissociates to give CuO on heating. It is difficult to produce it in the spinel structure as the $\text{CuO}-\text{Fe}_2\text{O}_3$ phase diagram indicates. Professor Yamaguchi suggests that the dissociation behavior of CuO from CuFe_2O_4 by heat can be summarized as follows (Yamaguchi, 1968).



For reasons mentioned here, it was expected that CuFe_2O_4 in $\text{NiCuZnFe}_2\text{O}_4$ also gives rise to this dissociation behavior. To prevent the compressive stresses on the ferrite grain boundaries and to have high-performance multilayer chip ferrite components, we forced CuFe_2O_4

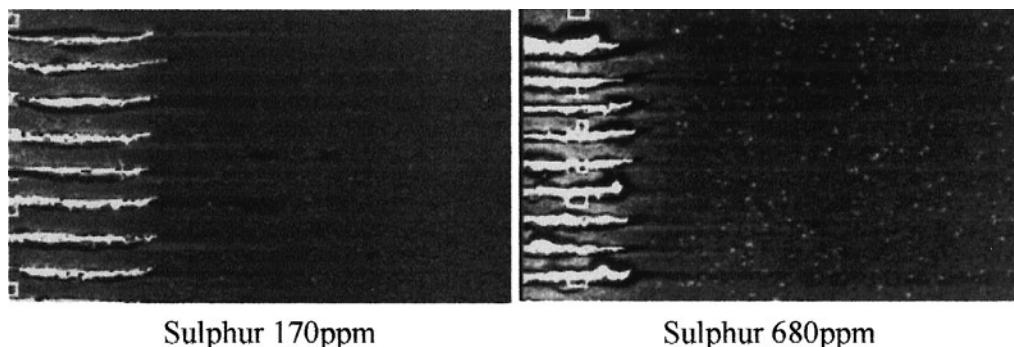


FIGURE 64 Effect of sulfur content on the Ag-diffusion behavior in multilayer chip inductor.

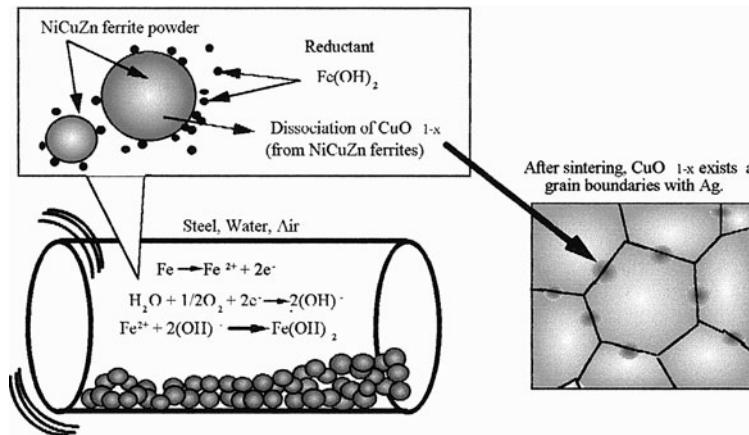


FIGURE 66 Mechanochemical reaction during the milling process.

into $(\text{NiCuZn})\text{Fe}_2\text{O}_4$. From in-depth investigations, we realized that controlling of the dissociation behavior of CuO from CuFe_2O_4 during the milling process is most important to prevent compressive stresses at the grain boundaries of multilayer chip inductor ferrites. These results may be better understood from the point of view of redox reactions during the milling. Figure 66 shows the mechanochemical reaction during the milling process. First, in the case of the conventional ferrites milling process, it is well known that $\text{Fe}(\text{OH})_2$ contamination is created in the steel milling-pot. The latter is a very strong reducing agent, which requires oxygen. When there is not enough oxygen in the milling-pot, it is thought that $\text{Fe}(\text{OH})_2$ takes oxygen from NiCuZn -ferrite rather than from air, and that $(\text{NiCuZn})\text{Fe}_2\text{O}_4$ changes to the segregated CuO_{1-x} and a spinel phase. Therefore, after sintering, CuO_{1-x} exists at the ferrite's grain boundaries, as was confirmed using the TEM/EDX observation (Nakano *et al.*, 1995). In order to obtain the high-performance multilayer chip ferrite components, chemical compositions of ferrite materials as well as controlling of stresses from the internal conductor and the grain boundaries are most important.

4. Hard Ferrites

Hard ferrites (ferrite magnets) are nowadays mostly used as permanent magnets all over the world. Among various hard magnetic materials, the production weight ratio of hard ferrites has been estimated around 95% since the 1980s, and the world market is expected to grow steadily. The reason for this dominance in mass production is due to unsurpassed cost performance and high chemical and magnetic stability.

J. J. Went *et al.* (1952) (Philips N. V.) systematically studied Ba-ferrites with hexagonal structures and they announced the M-type Ba-ferrite as a new permanent mag-

net material, "Ferroxdure." Since then, a great number of attempts have been made to improve the magnetic properties of these materials. Typical and successful examples are (1) the discovery of Sr-ferrite (Cochardt, 1963) having larger magnetic anisotropy (K_1); (2) the development of dry- and wet-pressing technology under a magnetic field (Fisher, 1983; Stuijts *et al.*, 1954), which doubles the residual flux density (B_r); and (3) addition of SiO_2 – CaO to control the microstructure. By using these technologies plus technologies for cost efficiency, hard ferrites have now become the main current for mass production. Recently, M-type Sr-ferrites with La and Co are found to have larger magnetization and anisotropy, and are starting to be produced as the highest-grade material.

a. Compositions and fundamental properties of hard ferrites. Figure 67 shows a Magnetoplumbite (M)-type crystal structure whose chemical formula is $\text{SrFe}^{3+}_{12}\text{O}_{19}$ (Kojima, 1982; Landolt-Bornstein). Oxygen ions, which form hexagonal closed packing, are partially substituted by strontium ions (Sr^{2+}) or barium ions (Ba^{2+}). Ferrous ions (Fe^{3+}) exit between the oxygen ions and their magnetic spin moments of 3d-electrons are connected with each other. In the M-structure, uniaxial magnetic anisotropy along the c -axis is generated, which is required for a permanent magnet. W- and X-type ferrites (Graetsoh *et al.*, 1984; Kojima *et al.*, 1985; Leccaube *et al.*, 1986, 1987) also having uniaxial anisotropy lie on the line between M and S (spinel) as shown in Fig. 68. There are many more related compounds than M-type ferrites for the reason that W- and X-type ferrites have divalent cation sites. Table XV shows intrinsic magnetic properties of some of these hexa-ferrites. The M-type Sr-ferrite (SrM) has a 10% larger anisotropy constant (K_1) than the M-type Ba-ferrite (BaM) (Shirk and Buessen, 1969). Magnetization (J_s) of Zn_2W - or Fe^{2+}W -ferrites is at least 10% larger than that

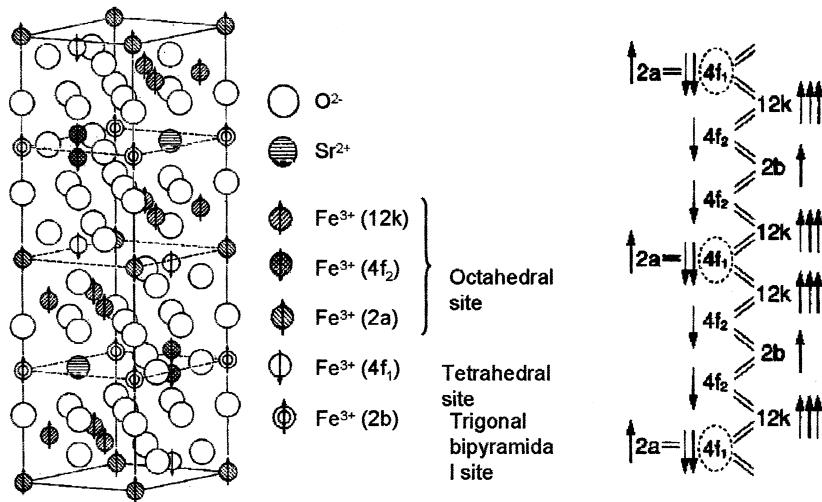


FIGURE 67 M-type Sr-ferrite. Crystal structure and exchange interaction scheme of magnetic moment of a ferrous ion (Fe^{3+}).

of M-type ferrites, although K_1 of Me^{2+}W -ferrites is considerably low except Fe^{2+}W .

There are a great number of studies concerning cation substitution for these hexa-ferrites (Albanese and Deriu, 1979; Mones and Banks, 1958). For a practical example, addition of Al_2O_3 is one of the conventional way to obtain higher H_{cJ} (but lower B_r). This is because Al^{3+} substitutes for Fe^{3+} in M-type ferrite ($\text{SrAl}^{3+}_{x}\text{Fe}^{3+}_{12-x}\text{O}_{19}$) and enhances the anisotropy field H_A ($=K_1/J_s$) (Haneda and Kojima, 1973). A large decrease of magnetization (J_s) is due to the fact that nonmagnetic ions Al^{3+} replace Fe^{3+} at 12 k sites, which have up-spin moments. Substitution of

two different ions is possible if the total valence is compensated ($\text{SrMe}^{2+}_{x/2}\text{Me}^{4+}_{x/2}\text{Fe}^{3+}_{12-x}\text{O}_{19}$) (Lotgering *et al.*, 1961). Fine powder of Ba-ferrite-containing Co and Ti (e.g., $x = 0.4\text{--}0.9$ in $\text{BaCo}_{x/2}\text{Ti}_{x/2}\text{Fe}_{12-x}\text{O}_{19}$), having reduced H_{cJ} (32–160 kA/m), is used for a magnetic recording medium (Kubo *et al.*, 1982).

Meanwhile, La^{3+} has the largest ionic radius among lanthanides (slightly smaller than Sr^{2+}) and has the largest solubility limit in M-type ferrite (Deschamps and Bertaut, 1958). There are two possible ways to compensate the valence difference between Sr^{2+} and La^{3+} . First is the combination with alkali metal ions ($\text{A}^+_{0.5}, \text{La}^{3+}_{0.5}, \text{Fe}_{12}\text{O}_{19}$,

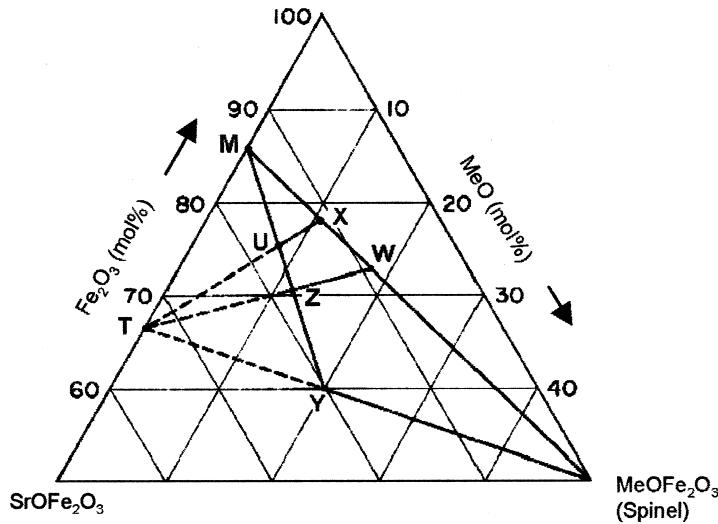


FIGURE 68 Chemical compositions of ferrimagnetic hexagonal compounds. Me represents a divalent ion among the first transition elements. SrOFe_2O_3 and T -phase do not exist as stable phases.

TABLE XV Fundamental Magnetic Properties of Hexa-Ferrite Magnet Materials

Symbol	Chemical formula	$\sigma s \times 10^{-6}$ (Hm ² /kg)	Tc (K)	H _A (MA/m)	K ₁ (J/m ³)	Reference
SrM	SrO·6Fe ₂ O ₃	1.14	733	1.59	0.35	Shirk and Buessen (1969)
SrFe2W	SrO·2FeO·8Fe ₂ O ₃	1.23	783	1.55		Kojima <i>et al.</i> (1985)
SrZn2W	SrO·2ZnO·8Fe ₂ O ₃	1.25	643			Graetsch <i>et al.</i> (1984)
SrZnX	2(SrO·ZnO·7Fe ₂ O ₃)	1.21	703	0.98	0.26	Leccabue <i>et al.</i> (1987)
BaM	BaO·6Fe ₂ O ₃	1.08	723	1.35	0.32	Shirk and Buessen (1969)
BaFe2W	BaO·2FeO·8Fe ₂ O ₃	1.19	773	1.4		Kojima <i>et al.</i> (1985)
BaZn2W	BaO·2ZnO·8Fe ₂ O ₃	1.23	643	1	0.26	Leccabue <i>et al.</i> (1986)

A = Na, K, Rb) (Summergrand and Banks, 1957). Second is the combination with divalent ions ($\text{Sr}^{2+}_{1-x}\text{La}^{3+}_x\text{ME}^{2+}_{12-x}\text{O}_{19}$). By means of this second way, Fe³⁺ can be replaced by 3d transition divalent ions only (e.g., Fe²⁺, Co²⁺, Ni²⁺, or Zn²⁺), to change the intrinsic magnetic properties. Some studies were reported from this point of view (Mulay and Sinha, 1970; van Diepen and Lotgering, 1974). Actually, at the composition of Sr_{0.7}La_{0.3}Fe_{11.7}Zn_{0.3}O₁₉, around 5% higher magnetization is obtained compared to conventional M-type Sr-ferrite, while retaining the K₁ value of conventional M-type Ba-ferrite (Taguchi *et al.*, 1997) as shown in Table XVI. Furthermore, at the composition of Sr_{0.7}La_{0.3}Fe_{11.7}Co_{0.3}O₁₉, K₁ is improved more than 15% compared to conventional M-type Sr-ferrite (Iida *et al.*, 1999) as also shown in Table XV that it is a remarkable value of K₁ = 4.2×10^5 J/m³. At this composition, saturation magnetization J_s is also improved around 2% at the same time resulting in high-performance magnets (Ohta, 1963). Divalent ions can be introduced by monovalent anion substitution. Studies on fluorine substitution in M-type ferrite ($\text{SrMe}^{2+}_x\text{Fe}^{3+}_{12-x}\text{O}_{19-x}\text{F}_x$) were reported from this point of view (Frei *et al.*, 1960; Robbins *et al.*, 1963).

b. Factors of magnetic properties. From a viewpoint of materials research, residual flux density (Br) and intrinsic coercivity (HcJ) are the most important properties. Both Br and HcJ consist of factors originated in the

crystal structure and micro- or nanostructure, as shown in Fig. 69. In order to improve the magnetic properties, all these factors should be improved simultaneously.

Br is determined by the product of desnsity, the degree of orientation, and the saturation magnetization (J_s). The M-type Sr-ferrite has a J_s value of about 0.465 T. The density and the degree of orientation have upper limits of about 98% in the case of sintered magnets, which afford the highest values. Therefore, the practical limit of Br in a sintered ferrite magnet with conventional M-type composition is around 0.445 T (=0.465 × 0.98 × 0.98), resulting in (BH) max of 37 kJ/m³.

Concerning HcJ, it is assumed that intrinsic coercivity in ferrite magnets originates in the magnetic behavior of “single domain particles” (Kitakami *et al.*, 1988; Meiklejohn, 1953). In general, magnetic materials form a “magnetic domain structure” in order to reduce static magnetic energy. Boundaries between two domains are magnetic walls where magnetic moments gradually change their direction. When reducing the particle or grain size of magnetic materials, the “nonmagnetic wall state (=single domain state)” becomes energetically more stable. Moreover, when this single domain state is realized, intrinsic coercivity will be maximized because magnetic reversal mechanisms are limited only to spin rotation. The maximum HcJ value corresponds to the anisotropic field (H_A = 2K₁/J_s), which is estimated at 1.5 MA/m for M-type Sr-ferrite. The critical diameter, below which a single

TABLE XVI Magnetic Properties of High-Performance M-Type Sr-Ferrite Containing Lanthanum

Composition	Fundamental properties			Sintered samples						
	J _s (T)	K ₁ (J/m ³)	Tc (K)	Br (T)	HcJ (kA/m)	Jr/J _s (%)	Hk/HcJ (%)	(BH) _{max} (kJ/m ³)	density (mg/m ³)	reference
Sr _{0.7} La _{0.3} Fe _{11.7} Zn _{0.3} O ₁₉	0.485	0.33	693	0.46	207	98.3	94.4	41	5.06	Taguchi <i>et al.</i> (1997)
Sr _{0.7} La _{0.3} Fe _{11.7} Co _{0.3} O ₁₉	0.470	0.42	714	0.445	383	97.8	83.0	39	5.05	Iida <i>et al.</i> (1999), Nishio <i>et al.</i> (1999)

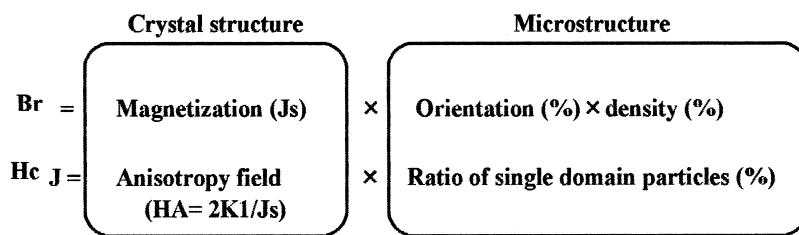


FIGURE 69 Factors of magnetic properties (Br and HcJ) in ferrite magnet.

domain is stable, is calculated at around $1\ \mu\text{m}$ for M-type Sr-ferrite assuming an isolated sphere.

c. Production process for sintered hard ferrites. Figure 70 shows the conventional manufacturing process of ferrite magnets. The raw materials strontium carbonate (or barium carbonate) and iron oxide are mixed and calcined at $1250\text{--}1300^\circ\text{C}$ in air. At this stage, ferrite formation must be concluded, especially in the case of anisotropic ferrite production, as it is necessary to orient the ferrite particles at the pressing stage. For this reason, higher calcination temperatures (above 1250°C) have been conventionally employed, causing grain growth. The cal-

cined products, which are usually hard pellets of more than a few millimeters in size, are crushed and pulverized. For the production of anisotropic magnets, calcined pellets are first dry-pulverized and then wet-milled to submicron order.

Ferrite magnets can be classified into two groups according to their manufacturing process: isotropic magnets and anisotropic magnets. Anisotropic magnets are mainly manufactured by pressing in a magnetic field of $0.4\text{--}1.2\ \text{MA/m}$. There are two methods for pressing: wet and dry. Wet-pressing is a process for compacting slurry of fine-milled powder mixed with a solvent such as water in a magnetic field. Although the flotation of the particles

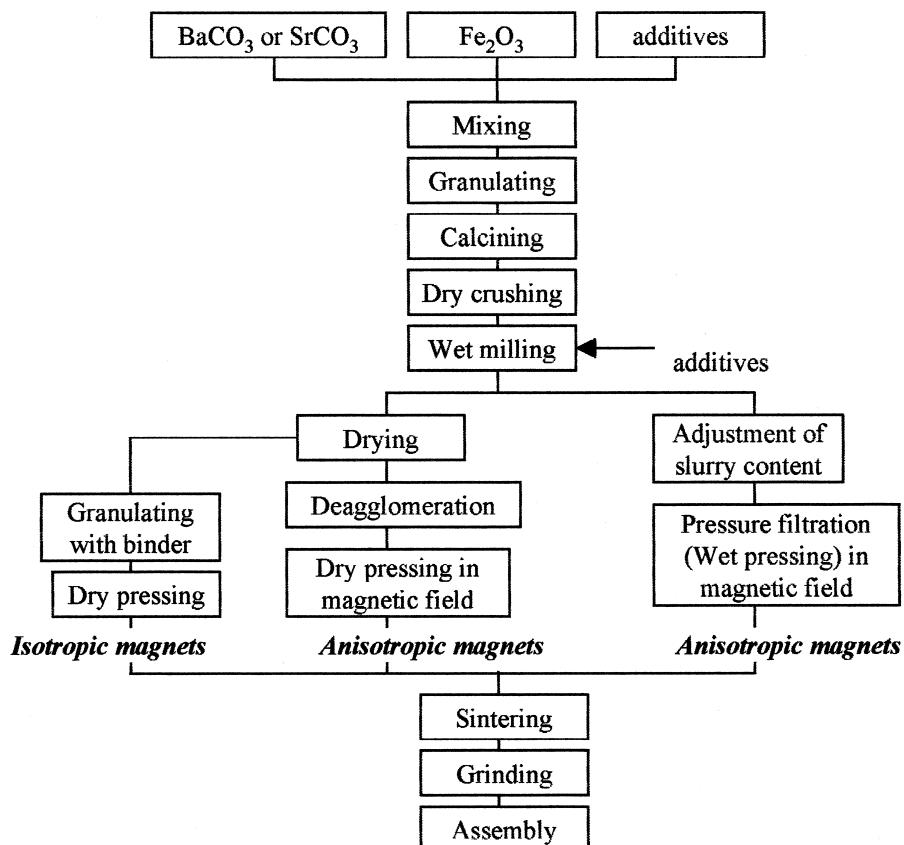


FIGURE 70 Production process of sintered ferrite magnets (hard ferrite).

in the solvent facilitates their rotation and orientation, dehydratation takes additional time. SiO_2 and CaCO_3 are usually added at the pulverizing stage in order to promote densification and inhibit grain growth at the next sintering stage. Compacts are sintered at 1200–1250°C in air. During the sintering, which is determined by the molar ratio $\text{Fe}_2\text{O}_3/\text{SrO}$, SiO_2 , CaO , and excess SrO form the glassy second phase at grain-boundaries (Kools, 1992). The ratio and amounts of these additives influence sintering behavior viz. final density, aspect ratio of the grains, and shrinkage ratio.

i. Process for submicron particles and grains. Large grains in excess of 1 μm show multidomain behavior and reduce the coercivity. In order to reach the theoretical value of $\text{HcJ} (= \text{H}_A)$, the ratio of single domain particles or grains should be increased. Therefore, control of grain size to under a micron is significant for achieving high HcJ as mentioned earlier. Taking into account grain growth, it is desirable that ferrite particles for sintered magnets should be around 0.3 μm . It is difficult to obtain this powder by conventional processing (e.g., simply long milling), because very fine particles (0.01–0.02 μm) also appear. These ultrafine particles have low magnetic properties due to thermal fluctuation, and lower the productivity of sintered compounds especially at the pressing stage. Similarly, particles prepared by the coprecipitation method, hydrothermal synthesis, or by glass crystallization are too small (under 0.1 μm) to be oriented in a magnetic field. Obtaining submicron-sized primary particles after calcination and before pulverization has an advantage for having an adequate size between 0.1 and 1 μm in the production of sintered magnets. From this point of view, calcination temperatures should be lowered by means of careful mixing of fine raw materials. In addition, the molar ratio $\text{Fe}_2\text{O}_3/\text{SrO}$ affects the grain growth at the calcination and final sintering. When the molar ratio is near stoichiometric ($\text{Fe}_2\text{O}_3/\text{SrO} = 6.0$) or Fe-rich (>6), grain growth is inhibited.

As the main raw material for ferrites, selection of iron oxide is very important. The points are the primary particle size, impurity level, and cost. Iron oxides from pickling liquors ($\alpha\text{-Fe}_2\text{O}_3$), having around a 0.2- to 0.7- μm primary particle, are usually used for high-grade materials. SrCO_3 should be dispersed well among the Fe_2O_3 particles, but the small particle size has less importance due to the self-decomposition during the calcination. Wet mixing process has an advantage for high dispersion. In that case, addition of dispersant and pH control of the slurry should be of concern. By means of this high-dispersion method of the fine raw materials (Fe_2O_3 and SrCO_3), the calcining temperature should be lowered to around 1200°C to avoid unnecessary grain growth. By the ideal process in the laboratory, the average size of the particles can be around 0.3 μm , having the magnetic properties with

$\sigma_s = 1.12 \times 10^{-6} \text{ Hm}^2/\text{kg}$ and $\text{HcJ} = 422 \text{ kA/m}$ for M-type Sr-ferrite (Taguchi *et al.*, 1992). Similar powders can be also obtained by the dry-mixing process.

ii. Process for high orientation. It is believed that single domain particles attract each other magnetically thereby creating agglomerates, which prevent orientation. Magnetic flux density at the working point, which is determined by the particle shape, relates to the magnetic agglomeration force. Thus, the surface flux density of the particle must be reduced by reducing HcB . It was found that dry-vibration pulverization in particular was effective in reducing the HcB values (Taguchi *et al.*, 1992).

The wet process using an organic solvent (e.g., toluene or xylene) and surfactant (e.g., stearic acid or oleic acid) made a remarkable improvement in the degree of orientation for submicron particles (Taguchi *et al.*, 1992). In the case of a magnetic fluid, for instance, it is well known that fine ferrite particles (0.01 μm) are dispersed well in organic solvents with added surfactant such as oleic acid. In the same way, Sr-ferrite particles coated with oleic acid or stearic acid dispersed well in toluene or xylene during the milling and pressing stages. This gave rise to the remarkable improvement of orientation under a magnetic field. In the case of water-based wet process, polymeric dispersant such as polycarboxylic acid-type compounds has been conventionally used. Recently low-molecular-type dispersant such as gluconic acid was found to be effective in improving the orientation of hard ferrite-submicron particles.

d. Development of high-performance hard ferrites. Figure 71 shows the progress of hard ferrites in the graph of Br and HcJ . First, processes for submicron-sized powder, for high orientation and for high desity, have been developed. Second, a new composition containing LaZn or LaCo was discovered to increase magnetization (J_s) and anisotropy (K_1). As a result, hard ferrites with extremely high values of $\text{Br} = 450 \text{ mT}$, $\text{HcJ} = 380 \text{ kA/m}$ or $\text{Br} = 0.46 \text{ T}$ (4.6 kG) and $(\text{BH})_{\text{max}} = 41 \text{ kJ/m}^3$ (5.2 MGOe) were obtained. Furthermore, the temperature coefficient of HcJ (100–400 K) was also improved viz. $\Delta \text{HcJ}/\Delta T = 1000 \rightarrow 410 \text{ A/m/deg}$ or $\Delta \text{HcJ}/\text{HcJ}/\Delta T = 0.33 \rightarrow 0.1\%/\text{deg}$ (Ohta, 1963). By the technology from micron order (control of particle- and grain-size distribution) to angstrom order (substitution by other elements), the magnetic properties of hard ferrites are still pushing back the theoretical limits.

5. Concluding Remarks

The demand for both soft and hard ferrites has been growing, and ferrites will expand markedly in both quantity and extent of application as the need for ferrites of higher quality increases. Raw materials, of which iron oxides are

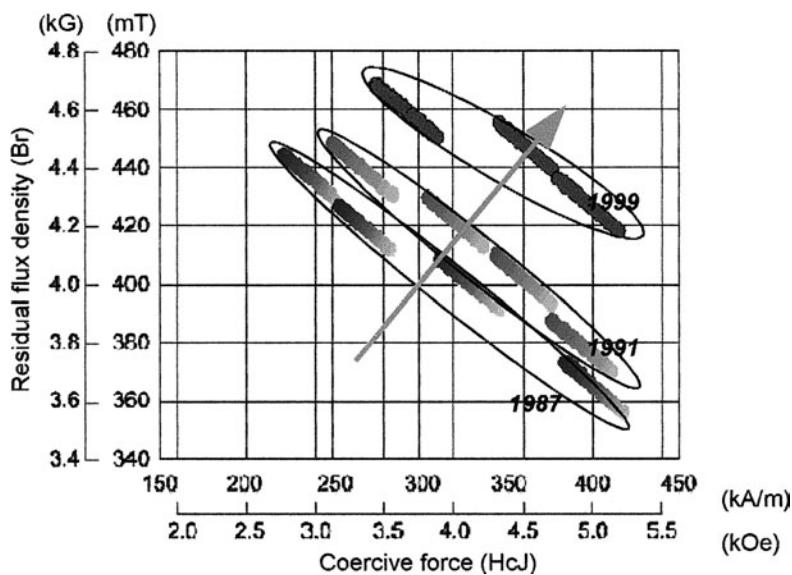


FIGURE 71 Progress of commercial grade hard ferrites compared by Br and HcJ. (Wet anisotropic grades are only shown.)

a major constituent, play a decisive role in improving the quality as well as lowering the cost of ferrites. By combinations of improved raw materials, compositional and processing improvements, and new classes of both soft and hard ferrites will be developed. In order to produce higher-performance soft ferrites, highly pure ferrite powder is a key. Small amounts of ingredients, as well as the homogeneity of ferrite powder and firing program, have a strong effect on microstructure, and hence the magnetic properties. Grain boundary chemistry is of prime importance not only for low loss characteristics but also for higher permeability.

Furthermore, new applications of ferrites such as EMC use and biochemical applications will be expanded. Multilayer technology will proceed more rapidly to keep up with the miniaturization of electronic devices. Ferrites, as the key materials of the electronics boom, will have fulfilled a major need of the electronics industry in the new millennium.

Bibliography for Section II.B can be found at the end of this article.

C. Porous Ceramics For Filtration

Tomonori Takahashi*

Filtration membranes made of porous ceramics are commercialized by several companies. The membranes are generally used to remove solid particles from liquids or gases, and membranes for gas separation are also being developed. The membranes are classified as follows:

*NGK Insulators, Ltd.

1. Shape: “tube,” “monolith,” “honeycomb,” and “plate”
2. Structure: “a symmetric membrane” which has homogenous microstructure, and “an asymmetric membrane” which consists of a substrate for mechanical supporting, and filtration or separation layer coated on the substrate
3. Materials: “porcelain,” “oxides” such as alumina, titania, zirconia and silica, silicon nitride, silicon carbide, carbon, etc.
4. Operation: “dead-end filtration” where the whole feed liquids and gases pass through a membrane; and “cross-flow filtration,” where a part of the feed liquids and gases pass through a membrane, and the feed mixture is concentrated

1. Microstructure

The portion which has filtration or separation function is a porous ceramic body or layer formed by partial sintering with ceramic grains with well-controlled grain size and its distribution. Figure 72 shows a typical microstructure of a porous alumina for the filtration. The bonding between the gains brings the mechanical strength. The spaces surrounded by the grains are pores through which the liquids and gases pass, and the solid particles are mechanically rejected if they are larger than the pores. As shown, the pore size is much related to the grain size. The minimum pore size is around $0.05 \mu\text{m}$ because the minimum size of the grains with well-controlled grain size and its distribution is around $0.1 \mu\text{m}$.

A porous ceramic layer with smaller pores is formed by appropriate heating of gel layer coated on a substrate by

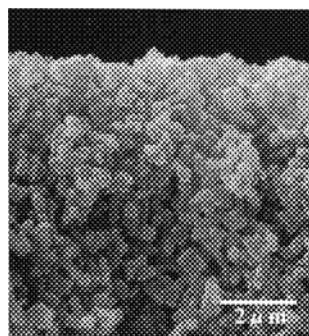


FIGURE 72 Porous alumina.

sol-gel process. The gel layer is converted to a layer consisting of very small grains and continuous pores among the grains. The minimum pore size is around 1 nm. **Figure 73** shows a microstructure of a porous layer with small titania grains prepared through sol-gel process. In addition these porous ceramic membranes are prepared by anodic oxidation, CVD and hydrothermal treatment.

Volume fraction of the pores is “porosity,” which is generally 20–70%.

2. Evaluation

Membranes are evaluated with permeation of liquids or gases without particles under a specific pressure difference, and a size of particles which can be rejected or trapped. A porous ceramic membrane with higher porosity and larger pore size shows higher permeation, while its mechanical strength is lower. The permeation is in reverse proportion to the membrane thickness. Since a substrate of an asymmetric membrane is usually thick for the mechanical supporting, its porosity and pore size have to be optimized considering permeation and mechanical strength.

Other than the mechanical rejection the solid particles are occasionally trapped by sorption on the grains even if they are smaller than that of the pores. The size of particles which can be rejected and/or trapped is indicated as a

size of particles, a specific fraction of which cannot pass through the membrane, which is obtained by interpolating with fractions of monosized particles with kinds of size rejected and/or trapped by the membrane. The particles are protein with a particular molecular weight, well-dispersed latex spheres with a same size, and so on. The pore size and its distribution are other essential specifications of membranes, since the rejection and trapping of the particles are much influenced by operation conditions. The pore size is determined by the mercury porosimetry, air-flow method, and vaporization-precipitation method.

3. Application

Filtration membranes are applied to liquid filtration such as biomass removal for food, bacteria removal in city water preparation, waste water treatment, circulating water filtration, and slurry concentration; and gas filtration such as dust removal from incineration exhaust, and carbon particles removal from diesel exhaust. In the liquid filtration the solid particles are often flocculated to get agglomerates larger than pores. Solid particles accumulating on the surface of and inside the membrane are removed by “Back-Flushing,” which is reverse flow operation. Scales remaining even after the back-flushing bring increase of pressure difference required to keep a constant permeation. In this case the membrane has to be washed with washing aids such as acid and/or base, or replaced.

Viscous flow of gases are prohibited by pores less than 100 nm, and specific gas molecules are selectively passed through membranes by Knudsen diffusion or surface diffusion along pore walls where they selectively adsorb gas molecules. Experimental results are reported with porous silica or zeolite membranes. Pervaporation, where specific components can vaporize from a liquid mixture through a membrane with selective affinity on its pore walls at the liquid–gas interface, is applied to separation water from alcohol–water mixture.

Bibliography for Section II.C can be found at the end of this article.

III. CERAMICS FOR BIOMEDICAL APPLICATIONS

Tadashi Kokubo*

Bone is a composite in which inorganic hydroxyapatite small crystallites are deposited on organic collagen fibers fabricated into a three-dimensional structure. The former occupies 70 wt% and the latter 30 wt%. Living cells

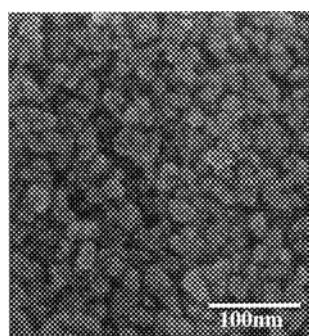


FIGURE 73 Porous titania.

*Department of Material Chemistry, Graduate School of Engineering, Kyoto University.

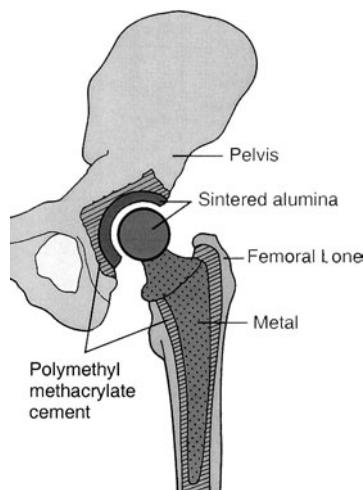


FIGURE 74 Total hip joint consisting of cup and head of sintered alumina, and stem of metal.

occupy only 1 vol%. Ceramics, which is defined as inorganic substance-based materials, therefore, can play important roles in repairing bone defects.

A. Ceramics for Joint Replacement

Conventional hip joint consists of a cup of ultra-high-molecular polyethylene, and head and stem of metals such as stainless steel, Co–Cr–Mo alloys, and titanium alloys. Wear debris of the polyethylene and metals induce the resorption of the surrounding bone, and hence loosening of the fixation of the joint. If the cap and head are replaced with a sintered high-density alumina, as shown in Fig. 74, production of the polyethylene and metal debris is eliminated. Rates of wear as well as increase in the friction coefficient of the joint are significantly decreased, since the sintered alumina has a high hardness, chemical durability, and hydrophilic property as well as high mechanical strength. These types of hip joints are already clinically used. Even when only the head is replaced with the sintered alumina, wear rate of the polyethylene cap is considerably decreased. This combination of the sintered alumina with the polyethylene is also clinically used in the hip joint as well as in the knee joint, as reported by Hulbert (1993).

Sintered yttria-stabilized zirconia has higher mechanical strength and fracture toughness than the sintered alumina. Therefore, the hip joint consisting of the head of the sintered zirconia and the cap of polyethylene is also clinically used.

B. Ceramics for Bone Replacement

Generally, artificial materials implanted into bone defects are encapsulated by a membrane of collagen fibers to be isolated from the surrounding bone. Some ceramics, how-

ever, come into direct contact with the surrounding bone without forming the collagen membrane around them and tightly bond to the bone. These ceramics are called bioactive ceramics. They and their composites with metals and organic polymers are clinically used as important bone-repairing materials.

1. Glasses

$\text{Na}_2\text{O}-\text{CaO}-\text{SiO}_2-\text{P}_2\text{O}_5$ glasses, called Bioglass,[®] form a bonelike apatite layer on their surfaces in the living body and bond to the living bone through the apatite layer. They are clinically used as periodontal fillers, maxillofacial implants, artificial middle ear bones etc., as reported by Hench (1998). Their mechanical strength is however, low.

2. Glass–Ceramics

Glass–ceramics precipitating crystalline apatite and β -wollastonite ($\text{CaO}\cdot\text{SiO}_2$), in a $\text{MgO}-\text{CaO}-\text{SiO}_2-\text{P}_2\text{O}_5$ glass, called Cerabone[®] A-W also forms a bonelike apatite layer on its surface in the living body and bond to the bone through the apatite layer. It has a higher bending strength than the human cortical bone and a higher bioactivity than the sintered hydroxyapatite, as reported by Kokubo (1993, 1998). It is clinically used as artificial vertebrae, intervertebral spacers, iliac spacers, bone fillers, etc. Figure 75 shows Cerabone[®] A-W which replaced a vertebra of a sheep and bonded to the surrounding cancellous bone.

3. Polycrystalline Ceramics

Sintered hydroxyapatite ($\text{Ca}_{10}(\text{PO}_4)_6(\text{OH})_2$) also forms a bonelike apatite layer on its surface in the living body and bonds to the bone. Its bending strength is higher than that of Bioglass[®] but lower than that of Cerabone[®] A-W. Its bioactivity is lower than those of Bioglass and Cerabone[®] A-W. They are clinically used as bone fillers in less loaded conditions, as described by LeGeros and LeGros (1993).

4. Calcium Phosphate Cements

When a powder mixture of tetracalcium phosphate ($\text{Ca}_4(\text{PO}_4)_2\text{O}$) and dicalcium phosphate anhydrous (CaHPO_4) is mixed with a sodium phosphate solution, the mixed paste hardens within 8 min to form a hydroxyapatite and bond to the living bone. A mixture of α -tricalcium phosphate ($\text{Ca}_3(\text{PO}_4)_2$), tetracalcium phosphate and dicalcium phosphate dihydrate ($\text{CaHPO}_4\cdot 2\text{H}_2\text{O}$) is mixed with a sodium succinic acid and a sodium chondroitin sulfate solution. The mixed paste also shows similar properties. They are clinically used as injectable bone fillers, as described by Chow (1998).

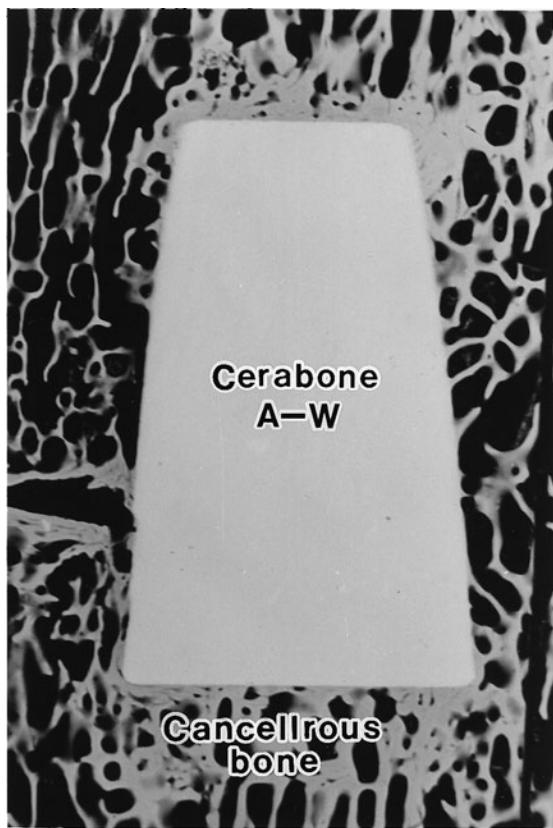


FIGURE 75 Cerabone® A-W which replaced vertebra of sheep.

5. Ceramic-Coated Metals

Titanium metal and its alloys coated with hydroxyapatite by the plasma spray method are clinically used as stem of the hip joint and artificial tooth root etc., because of its high fracture toughness and bioactivity. In this technique, however, hydroxyapatite powders are momentarily heated above 10,000°C, to be partially decomposed and molten. As a result, the coated layer is not stable for a long period in the living body. Recently, it was shown that titanium metal and its alloys spontaneously form a bonelike apatite layer on their surfaces in the living body and tightly bond to the bone, when they were previously subjected to NaOH solution and subsequent heat treatments to form a sodium titanate layer with a graded structure on their surfaces, as described by Kokubo (1993, 1998).

6. Ceramic–Polymer Composites

Bioactive materials described earlier have a higher elastic modulus than that of the human cortical bone. Bone repairing materials are desired to be compatible to the surrounding bone not only biologically but also mechanically. Polyethylene dispersed with hydroxyapatite powders in 40 vol% show low Young's modulus, ductility as well as bioactivity. It is clinically used as middle ear bone

etc., as described by Bonfield (1993). Its bioactivity, however, is not so high.

Preparation of composites in which apatite small crystallites are deposited on organic fibers fabricated into a three-dimensional structure similar to that of the collagen fibers in the bone is attempted by a biomimetic process. The resultant products are expected to exhibit analogous mechanical and biological properties to those of the natural bone, as described by Kokubo (1993, 1998).

Bibliography for Section III.A can be found at the end of this article.

SEE ALSO THE FOLLOWING ARTICLES

BIOMATERIALS, SYNTHESIS, FABRICATION, AND APPLICATIONS • **CERAMICS, CHEMICAL PROCESSING OF** • **COMPOSITE MATERIALS** • **GLASS-CERAMICS** • **MANUFACTURING PROCESS TECHNOLOGY**

BIBLIOGRAPHY

Bibliography for Section I.D

- Alberta, E. F., and Bhalla, A. S. (1998). *J. Korean Phys. Soc.* **32**, S1265–1268.
 Choi, S. W., Shrout, T. R., Jang, S. J., and Balla, A. S. (1989). *Ferroelectrics* **100**, 29–38.
 Galasso, F. (1969). "Structure, Properties and Preparation of Perovskite Type Compounds," Pergamon Press, New York.
 Harada, K., Shimanuki, S., Kobayashi, T., Saitoh, S., and Yamashita, Y. (1998). *J. Am. Ceram. Soc.* **81**, 2785–2788.
 Hosono, Y., Harada, K., Yamashita, Y., Dong, M., and Ye, Z. G. (2000). *Jpn. J. Appl. Phys.* **39**(9B).
 Jaffe, B., Roth, R. S., and Marzullo, S. (1955). *J. Res. Natl. Bur. Standard* **55**, November, 239–254.
 Kobayashi, T., Shimanuki, S., Saitoh, S., and Yamashita, Y. (1997). *Jpn. J. Appl. Phys.* **36**, 6035–6039.
 Kuwata, J., Uchino, K., and Nomura, S. (1982). *Jpn. J. Appl. Phys.* **21**, 1298–1302.
 Ouchi, H., Nagano, K., and Hayakawa, S. (1965). *J. Am. Ceram. Soc.* **48**, 630–635.
 Park, S.-E., and Shrout, T. R. (1997a). *Mater. Res. Innovat.* **1**, 20–25.
 Park, S.-E., and Shrout, T. R. (1997b). *J. Appl. Phys.* **82**, 1804–1809.
 Saitoh, S., Takeuchi, T., Kobayashi, T., Harada, K., Shimanuki, S., and Yamashita, Y. (1999). *Jpn. J. Appl. Phys.* **38**, 3380–3384.
 Shannon, R. O. (1976). *Acta Crystallogr.* **A32**, 751–767.
 Shimanuki, S., Saitoh, S., and Yamashita, Y. (1998). *Jpn. J. Appl. Phys.* **37**, 3382–3385.
 Shrout, T., Chang, Z. P., Kim, N., and Markgraf, S. (1990). *Ferroelectric Lett.* **12**, 63–69.
 Takeshima, Y., Shiratuya, K., Takagi, H., and Tomono, K. (1995). *Jpn. J. Appl. Phys.* **34**, 5083–5085.
 Tennery, V. J., Hang, K. W., and Novak, R. E. (1968). *J. Am. Ceram. Soc.* **51**, 671–674.
 Tsuzuki, E., Sakata, K., Ohara, G., and Wada, M. (1973). *Jpn. J. Appl. Phys.* **12**(10), 1500–1503.
 Yamamoto, T., and Ohashi, S. (1995). *Jpn. J. Appl. Phys.* **34**(9B), 5349–5353.

- Yamamoto, T., and Yamashita, Y. (1996). *Proc. ISAF'96*, East Brunswick, NJ, Aug. 1996. IEEE Catalog No. 96CH35948, 2, pp. 101–105.
- Yamashita, Y., and Shimanuki, S. (1996a). *Mater. Res. Bull.* **31**(7), 887–895.
- Yamashita, Y., Harada, K., Tao, T., and Ichinose, N. (1996b). *Integrated Ferroelectrics* **13**, 9–16.
- Yamashita, Y., Hosono, Y., Harada, K., and Ichinose, N. (2000). *Jpn. J. Appl. Phys.* **39**(9B).
- Yamashita, Y. (1994a). *Jpn. J. Appl. Phys.* **33**, 4652–4656.
- Yamashita, Y. (1994b). *Jpn. J. Appl. Phys.* **33**(9B), 5328–5331.
- Yamashita, Y., Harada, K., Hosono, Y., Natsume, S., and Ichinose, N. (1998). *Jpn. J. Appl. Phys.* **37**, 5288–5291.
- Yasuda, N., Ohwa, H., Kume, M., and Yamashita, Y. (2000). *Jpn. J. Appl. Phys.* **39**, L66–68.

Bibliography for Section I.H

- Matsuoka, M., Masuyama, T., and Iida, Y. (1971). “Nonlinear electrical properties of Zinc oxide ceramics,” *Jpn. J. Appl. Phys. Suppl.* **39**, 94–101.
- Yamaoka, N., Masuyama, M., and Fukui, M. (1983). SrTiO₃-based boundary layer capacitor having varistor characteristics, *Am. Ceram. Soc. Bull.* **62**(6), 698–700.

Bibliography for Section II.B

- Albanese, G., and Deriu, A. (1979). “Magnetic properties of Al, Ga, Sc, in substituted Barium ferrites: A comparative analysis,” *Ceramurgia Int.* **5**, 3.
- Akashi, M. (1961). *J. Phys. Soc. Japan* **30**, 708.
- Blank, J. M. (1961). *J. Appl. Phys.* **32**, 378S.
- Cochardt, A. (1963). *J. Appl. Phys.* **34**, 1273.
- Deschamps, A., and Bertaut, F. (1958). “Sur la substitution de baryum par une terre rare dans l’hexaferrite BaO₆Fe₂O₃,” *Acad. Sci.* **17**, 3069.
- Fisher, E. (1983). “Compacting unit for ferrite permanent magnets,” *Power Metallurgy Int.* **15**, 42.
- Frei, E. H., Schierber, M., and Shtrikman, S. (1960). “Hexagonal ferrimagnetic compound containing fluorine,” *Phys. Rev.* **118**, 657.
- Franken, P. E., and Stacy, W. T. (1980). “Examination of grain boundaries of MnZn ferrites by AES and TEM,” *J. Am. Ceram. Soc.* **63**, 315.
- Graetsch, H., Haberey, F., Leckebusch, R., Rosenberg, M. S., and Sahl, K. (1984). “Crystallographic and magnetic investigation on W-type-hexaferrite single crystals in the solid solution series SrZn_{2-x}Co_xFe₁₆O₂₇,” *IEEE Trans. Magn.* **MAG-20**, 495.
- Haneda, K., and Kojima, H. (1973). “Intrinsic coercivity of substituted BaFe₁₂O₁₉,” *Jpn. J. Appl. Phys.* **12**, 355.
- Iida, K., Minachi, Y., Masuzawa, K., Kawakami, M., Nishio, H., and Taguchi, H. (1999). “High performance ferrite magnets: M-type Sr-ferrite containing lanthanum and cobalt,” *J. Magn. Soc. Japan* **23**, 1093.
- Kanagawa, Y., Suzuki, T., Watanabe, H., and Nomura, T. (1988). “Multilayer Ferrite Surface Mount Device,” 8th Takei meeting 1.
- Kojima, H. (1982). “Ferromagnetic Materials 3,” pp. 305–391, North-Holland Publishing Company, Amsterdam.
- Kitakami, O., Goto, K., and Sakurai, T. (1988). “A study of the magnetic domains of isolated fine particles of Ba ferrite,” *Jpn. J. Appl. Phys.* **27**, 2274.
- Kojima, H., Miyakawa, C., Sato, T., and Goto, K. (1985). “Magnetic properties of W-type hexaferrite powders,” *Jpn. J. Appl. Phys.* **24**, 51.
- Kools, F. (1985). “The action of a silica additive during sintering of strontium hexaferrite,” *Sci. Sintering* **17**, 49.

- Kubo, O., Ido, T., and Yokoyama, H. (1982). “Properties of Ba ferrite particles for perpendicular magnetic recording media,” *IEEE. Trans. Magn.* **MAG-18**, 1122.
- Landolt-Bornstein: Neue Serie III/4b, “7 Hexagonal ferrites.”
- Leccabue, F., Panizzieri, R., Bocelli, G., Calestani, G., Rizzoli, C., and Suarez Almodovar, N. (1987). “Crystal structure and magnetic characterization of Sr₂Zn₂Fe₂₈O₄₆ (SrZn-X) hexaferrite single crystal,” *J. Magn. Magn. Mat.* **68**, 365.
- Leccabue, F., Panizzieri, R., Salviati, G., Albanese, G., and Sanchez Liamazares, J. L. (1986). “Magnetic and, morphological study of BaZn₂Fe₁₆O₂₇ hexagonal ferrite prepared by chemical coprecipitation method,” *J. Appl. Phys.* **59**, 2114.
- Lotgering, F. K., Enz, U., and Smit, J. (1961). “Influence of Co²⁺ ions on the magnetic anisotropy of ferrimagnetic oxides having hexagonal crystal structures,” *Philips Res. Repts.* **16**, 441.
- Meiklejohn, W. H. (1953). “Experimental study of the coercive force of fine particles,” *Rev. Mod. Phys.* **25**, 302.
- Mones, A. H., and Banks, E. (1958). “Cation substitution in BaFe₁₂O₁₉,” *J. Phys. Chem. Solids* **4**, 217.
- Mulay, V. N., and Sinha, A. P. B. (1970). “Synthesis and properties of some new ferrites of formula La³⁺Me²⁺Fe³⁺₁₁O₁₉,” *Indian J. Pure Appl. Phys.* **8**, 412.
- Nakano, A., Kanagawa, Y., Suzuki, T., and Nomura, T. (1990). “Multilayer Ferrite Surface Mount Device,” 10th Takei meeting 1.
- Nakano, A., and Nomura, T. (1999). “Multilayer chip inductors,” *Proc. of A. Cer. S.* **97**, 285.
- Nakano, A., Momoi, H., and Nomura, T. (1992). “Nano-Structure Control of NiCuZn Ferrites for Multilayer Chip Components,” *J. Jpn. Soc. of Power and Power Metallurgy* **39**, 612.
- Nakano, A., Momoi, H., and Nomura, T. (1992). “Effect of age on the microstructure of the low temperature sintered NiCuZn ferrites,” *Proc. Int. Conf. Ferrites* 1225.
- Nakano, A., and Nakah, I. (2000). “Low temperature sintering MgCuZn ferrites for multilayer chip components,” *Progressing Proc. of A. Cer. S.*
- Nakano, A., Sato, T., and Nomura, T. (1995). “Microstructure control of NiCuZn ferrites for multilayer ferrite chip components,” *Proc. of A. Cer. S.* **47**, 241.
- Nomura, T., Ochiai, T., Kitagawa, T., and Okutani, K. (1982). “Sintering of MnZn Ferrites for Powder Electronics,” *Am. Ceram. Soc.*, Fall meeting, pp. 57-BE–82F.
- Nishio, H., Iida, K., Minachi, Y., Masuzawa, K., Kawakami, M., and Taguchi, H. (1999). “Magnetocrystalline anisotropy of M-type Sr-ferrite containing lanthanum and cobalt,” *J. Magn. Soc. Japan* **23**, 1093.
- Nomura, T. (1992). *Proceedings of ICF* 6, 65.
- Nomura, T., and Takaya, M. (1987). “Multilayer ferrite surface mount device,” *Hybrids* **3**, 15.
- Nomura, T., and Nakano, A. (1992). “New evolution of ferrite for multilayer chip components,” *Proc. Int. Conf. Ferrites* 1198.
- Ochiai, T., and Okutani, K. (1985). *Proceedings of ICF* 4, 447.
- Ohta, K. (1963). *J. Phys. Soc. Japan* **18**, 685.
- Otsuki, E. (1992). *Proceedings of ICE* 6, 59.
- Paulus, M. (1966). “Properties of grain boundaries in spinel ferrites,” *Mat. Sci. Res.* **3**, 31.
- Roess, E., and Moser, E. (1961). *Z. Angew Phys.* **13**, 247.
- Roess, E., Hanke, I., and Moser, E. (1964). *Z. Angew Phys.* **17**, 504.
- Robbins, M., Lerner, S., and Banks, E. (1963). “Some effects of charge-compensating ions on cation site preferences in complex oxides,” *J. Phys. Chem. Solids* **24**, 759.
- Shirk, B. T., and Buessen, W. R. (1969). “Temperature dependence of Ms and K₁ of BeFe₁₂O₁₉ and SrFe₁₂O₁₉ single crystals,” *J. Appl. Phys.* **40**, 1294.

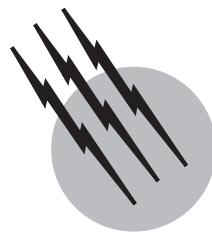
- Smit, J., and Wijn, H. P. (1959). "Ferrite," *Philips Res. Lab.*
- Sato, T., and Otsuki, E. (1996). *Powder Powder Metallurgy* **43**, 1393.
- Stuijts, A. L., Rathenau, G. W., and Weber, G. H. (1954). "Ferroxdure II and III, anisotropic permanent magnet materials," *Philips Techn. Rev.* **16**, 141.
- Summergrand, R. N., and Banks, E. (1957). "New hexagonal ferrimagnetic oxides," *J. Phys. Chem. Solids* **2**, 312.
- Taguchi, H., Hirata, F., Takeishi, T., and Mori, T. (1992). "High performance ferrite magnet," *Ferrites* (Proc. 6th international conference on ferrites), 1118.
- Taguchi, H., Takeishi, T., Suwa, K., Masuzawa, K., and Minachi, Y. (1997). "High energy ferrite magnets," *J. Phys. IV France* **7** (ICF7 Proc. 7th international conference on ferrites) C1–311.
- Takaya, M. (1988). "Multilayer chip inductor," *Proc. Int. Symp. Microelectronics*, Japan, 25.
- Takaya, M., and Nomura, T. (1994). "Chip inductor," *Electro Ceramics* **25**, 12.
- van Diepen, A. M., and Lotgering, F. K. (1974). "Mossbauer effect in LaFe₁₂O₁₉," *J. Phys. Chem. Solids* **35**, 1641.
- Went, J. J., Rathenau, G. W., Gorter, E. W., and van Oosterhout, G. W. (1952). "Ferroxdure, A class of new permanent magnet materials," *Philips Tech. Rev.* **13**, 194.
- Yamaguchi, T., (1968). *J. Jpn. Soc. Power Metallurgy* **15**, 445.
- Hsieh, H. P. (1996). Traditional liquid-phase separation application, and gas-phase and non-traditional separation applications. In "Inorganic Membranes for Separation and Reaction," Elsevier, Holland. JP96-133857A.
- Kita, H., Harada, T., Asamura, H., Tanaka, K., and Okamoto, K. (1998). "Conventional vs. Microwave Hydrothermal Synthesis of zeolite Membranes and their Pervaporation Properties," **ICIM-5**, 318–321.
- Kusakabe, K., and Morooka, S. (1998). The use of zeolite membranes for the separation of CO₂. *Proceedings of International Workshop on Zeolitic Membranes and Films*, 29–32.
- Mulder, M. (1991). Characterization of membranes. In "Basic Principle of Membrane Technology," Kluwer Academic Publishers, New York.

Bibliography for Section III.A

Bibliography for Section II.C

Asaeda, M. (1998). "Separation of Gaseous Mixtures by Inorganic Porous Membranes Prepared by Sol-Gel Process," **ICIM-5**, 2–5.

- Bonfield, W. (1993). In "An Introduction to Bioceramics" (L. L. Hench and J. Wilson, eds.), World Scientific, pp. 299–303, Singapore.
- Chow, L. C. In "Bioceramics," Vol. 11 (R. Z. LeGeros and J. P. LeGeros, eds.), pp. 45–49, World Scientific, Singapore.
- Hench, L. L. (1998). *J. Am. Ceram. Soc.* **81**, 1705–1728.
- Hulbert, S. F. (1993). In "An Introduction to Bioceramics" (L. L. Hench and J. Wilson, eds.), pp. 25–40, World Scientific, Singapore.
- Kokubo, T. (1998). *Acta Mater.* **46**, 2519–2527.
- Kokubo, T. (1993). In "An Introduction to Bioceramics" (L. L. Hench and J. Wilson, eds.), pp. 75–88, World, Scientific, Singapore.
- LeGeros, U. M., and LeGeros, J. P. (1993). In "An Introduction to Bioceramics" (L. L. Hench and J. Wilson, eds.), pp. 139–180. World Scientific, Singapore.



Ceramics, Chemical Processing of

P. Hubert Mutin

Bruno Boury

University of Montpellier, France

- I. Principles
- II. Precursor Synthesis
- III. Cross-Linking and Pyrolysis
- IV. Crystallization
- V. Precursor Processing and Applications
- VI. Outlook

GLOSSARY

Ceramic Hard, brittle, heat-resistant, and corrosion-resistant nonmetallic inorganic material.

Composite A complex material in which distinct, structurally complementary substances (metals, ceramics, or polymers) combine to produce structural or functional properties not present in any individual component.

Organoelement compound Compound derived from a hydrocarbon by substitution of a carbon atom by an element E (E = Si, B, P, Al, Ti, . . .); by extension, any organic compound containing one (or more) element E and C, N, O, or H atoms.

Polymeric precursor Polymer or oligomer (short polymer) that can be converted into a ceramic by pyrolysis.

Pyrolysis Heat-treatment in a controlled atmosphere, inert or reactive.

Sintering A process that transforms agglomerated powders into a coherent mass by heating without melting.

HIGH-PERFORMANCE CERAMICS are generally manufactured at high temperature by sintering of ceramic powders. The infusibility and insolubility of these powders limit this process to the preparation of bulk ceramic products. The pyrolysis of a synthetic organoelement polymer is an attractive alternative route for the preparation of ceramics. The main advantage of this route over conventional powder processing is the possibility of preparing a variety of ceramic shapes: fibers, monoliths, films, or powders; furthermore, the lower processing temperatures allow the synthesis of metastable ceramics. The successful

processing of carbon fibers and bodies by pyrolysis of organic polymers is based on the same concept—that is, shaping of the polymeric precursor followed by its conversion into an inorganic residue by pyrolysis. Since the preparation of ceramic fibers from organosilicon precursors by Verbeeck *et al.* and Yajima *et al.* in the 1970s, a wealth of organoelement precursors to nonoxide ceramics has been developed. This has required a multidisciplinary approach involving synthetic organic, inorganic, and polymer chemists as well as solid-state and materials scientists.

I. PRINCIPLES

The preparation of a ceramic product from an organoelement polymer usually involves five steps (Fig. 1):

- Synthesis of the precursor by polymerization of suitable organoelement monomers

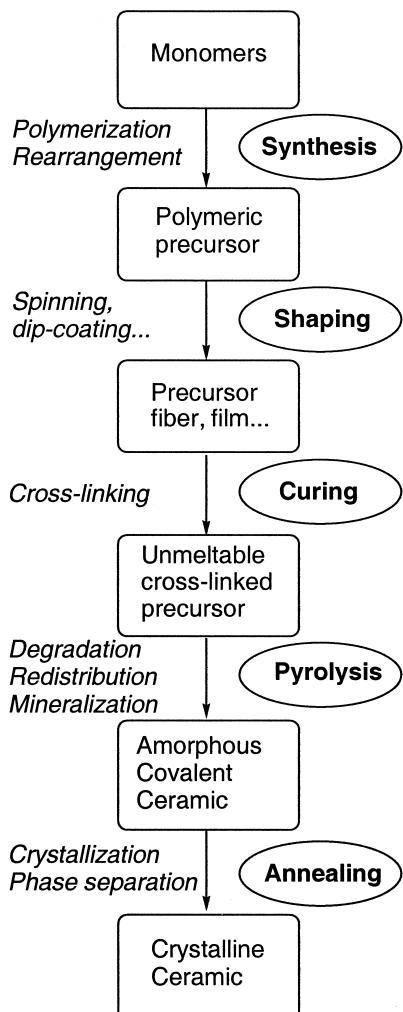


FIGURE 1 Flow diagram of ceramic processing by precursor pyrolysis.

- Shaping of the polymeric precursor
- Curing (or cross-linking) of the shaped precursor to form an unmeltable preceramic network
- Conversion of the organoelement network into an inorganic amorphous residue by pyrolysis in a controlled atmosphere
- Annealing of the amorphous ceramic product to convert it (if need be) into a crystalline product

Using this process, a wide variety of nonoxide ceramic products with different shapes, compositions, and structures can be prepared. In addition, metastable amorphous ceramics that are homogeneous on an atomic scale can be obtained, which lead, on crystallization, to controllable nano- or microstructures.

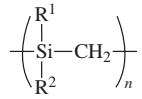
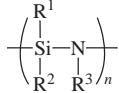
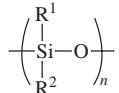
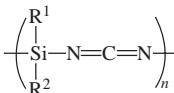
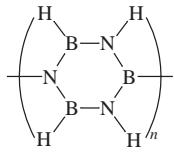
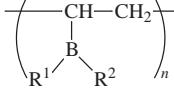
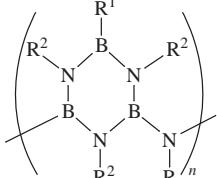
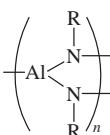
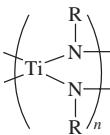
The formation of oxide ceramics will not be considered in this article as it is more related to the sol-gel process. The boundary between the precursor pyrolysis route and the sol-gel route is sometimes blurred. Actually, in the precursor pyrolysis process substantial amounts of the bonds in the final ceramic are formed during the pyrolysis step, while the organoelement precursor is converted into an inorganic residue. On the other hand, in the sol-gel process the oxide bonds are predominantly formed during the polymerization step, which results in an essentially inorganic network.

The ceramic composition first depends on the precursor composition, but it can be modified to a large extent by the curing and pyrolysis conditions. For instance, a polycarbosilane precursor containing Si, C, and H atoms can be converted into a Si–C (nonoxidative cure, inert pyrolysis), a Si–C–O (oxidative cure, inert pyrolysis), or a Si–C–N–O (oxidative cure, ammonia pyrolysis) ceramic.

Chemically stable, ionic-covalent bonds are required to form the backbone of the polymer; therefore, compared to the high number of known ceramics, only a few can be obtained from polymeric precursors due to the difficulty of polymer synthesis. In most cases, they are silicon-based ceramics and the other examples are limited to ceramics with boron, aluminium, or transition metals such as titanium or zirconium. The interest in silicon results from both the desirable properties of silicon-based ceramics (e.g., SiC and Si₃N₄) and the accessibility of silicon-based polymers.

Organosilicon polymers or copolymers with Si–Si, Si–C, Si–N, or Si–O bonds in their backbones lead after pyrolysis to various amorphous or crystalline silicon-based ceramics in the Si–C–N–O quaternary system (Table I). Organosilicon precursors containing different elements E (e.g., boron, titanium, or aluminium) can also be obtained, leading to ceramics in the Si–E–C–N–(O) system. Silicon-free polymers with B–N, B–C, or M–N bonding (M = Al, Ti) have also been prepared, leading to ceramics in the E–C–N system.

TABLE I Basic Organoelement Polymers Precursors to Ceramics

Nomenclature	Example	Ceramic
Polysilanes		SiC ^a Si ₃ N ₄ ^b
Polycarbosilanes		SiC ^a Si ₃ N ₄ ^b
Polysilazanes		Si-C-N ^a Si ₃ N ₄ ^b
Polysiloxanes		Si-O-C ^a Si-O-N ^a
Polysilylcarbodiimides		Si-C-N ^a Si ₃ N ₄ ^a
Polyborazlenes		BN
Polyvinylboranes		B-N-C ^a BN ^b
Polyborazinylamines		BN ^b
Polyaluminumimides		TiN ^b
Polytitaniumimides		AlN ^b

^a Inert atmosphere.^b NH₃ atmosphere.

II. PRECURSOR SYNTHESIS

A. Design of a Good Preceramic Polymer

A “good precursor” is an inexpensive, easy to handle, polymer, with a composition and a rheology allowing its

conversion in high yield to a ceramic with the desired shape and composition. This crude definition allows one to emphasize that the properties of the precursor must be adjusted for each application. In addition, to decrease the cost of the final ceramic products, the precursor should be inexpensive and therefore should be readily synthesized from commercially available, low-cost monomers.

1. Ceramic Composition

The ceramic composition first depends on the composition of the precursor, which supplies elements such as Si, B, C, N, etc. There is a significant modification of the composition during the pyrolysis. Depending on the pyrolysis conditions, several elements can be removed partially (C, N, B, Si, . . .) or completely (H, C, N). On the other hand, some elements (O, B, N, . . .) can be introduced during the curing step or pyrolysis using a reactive gas such as NH₃ or O₂. The pyrolysis in an inert atmosphere often results in ceramics containing an excess of carbon, leading to the formation of a “free-carbon” phase.

To adjust the composition of the final ceramic, it is often necessary to tune the composition of the precursor by design of the monomer, copolymerization of different monomers, or chemical modification of the polymer. Mixtures of polymers can, in principle, be used but this often leads to miscibility or mixing problems.

2. Ceramic Yield

The conversion by pyrolysis in an inert atmosphere of an organoelement precursor into an inorganic ceramic takes place with loss of gaseous by-products. To maximize the ceramic yield, the elemental composition of the precursor should be, in principle, as close as possible to that of the target ceramic. For instance, let us consider the ideal conversion into pure silicon carbide of precursors with elemental compositions SiCH₄, SiC₂H₆, and SiC₁₂H₁₀; the theoretical SiC yields for these precursors are 90.1, 69.0, and 22.0%, respectively. If the ceramic residue contains excess carbon, it should be taken into account to calculate the theoretical ceramic yield.

Actually, ceramic yields lower than the theoretical ones are often observed, due to two processes:

1. Distillation of low-molecular-weight oligomers
2. Thermal decomposition of the polymer into volatile organoelement compounds

Both problems can be solved by cross-linking the polymer chains, leading to a tridimensional network (see Fig. 2); however, this cross-linking completely modifies the rheology of the precursor, leading to an unsoluble, unmeltable solid. Consequently, the cross-linking is usually

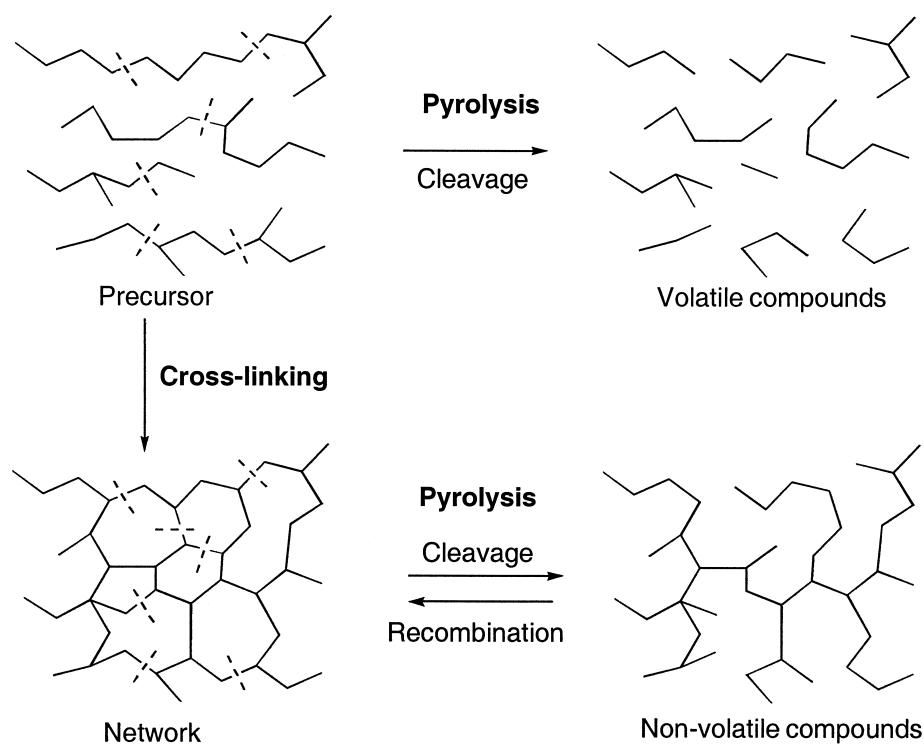


FIGURE 2 Cross-linking of polymeric precursors.

performed after the shaping of the precursor, during the curing step. This in turn implies that the precursor has many reactive substituents that can react with each other or with an additive to connect the chains.

3. Rheology of the Precursor

The rheological requirements are very different depending on the shaping of the ceramic envisaged, and the precursor can be, for example, a meltable or soluble solid or a low-viscosity liquid (see Fig. 6). Volatile oligomers can be used for chemical vapor deposition (CVD), another chemical route to ceramics. The physical properties of a precursor, such as melting point (crystalline polymer) or glass transition temperature (amorphous polymer), depend on numerous factors such as the nature of the bonds in the polymer backbone, the bulkiness of substituents, and the molecular weight and structure of the polymer (linear, branched, or cross-linked). The viscosity in the liquid state is an important parameter. Viscosity may be adjusted by addition of solvent or by controlling the molecular weight of the precursor. Low-molecular-weight polymers with low melt viscosities can be prepared using monofunctional monomers (chain-capping). Conversely, the viscosity may be increased by increasing the molecular weight of the precursor, for instance by using tri or tetrafunctional monomers.

4. Stability

To ensure a stable viscosity and reproducible shaping, the preceramic polymers must be stable at the processing temperature. In addition, even if chemists know how to handle highly reactive products, stability to air and moisture as well as thermal stability will always be advantageous since they make the polymer much easier (and less expensive) to process. These properties are also important for the “pot-life” of the preceramic polymer—that is, the time of storage and the storage conditions. Preservatives may be used to decrease the oxidizability of some polymers.

B. Precursor Synthesis

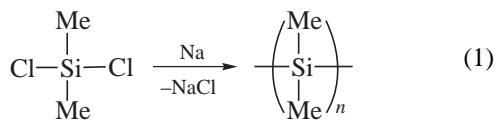
We have seen that the preparation of a good preceramic polymer requires the simultaneous control of its composition, reactivity, and structure. This can be done by design of the monomer, choice of the polymerization reaction, or copolymerization of different monomers with different substituents and functionalities. The choice of the substituents attached to the backbone is of the utmost importance. Reactive substituents allow the cross-linking of the precursor (during the synthesis or the curing step), which is essential for obtaining high ceramic yields. Si–H, (B–H), vinyl, and N–H groups are the most common reactive substituents found in preceramic precursors. Reactive

substituents also permit one to modify the composition of the polymer, in order to introduce additional elements into the final ceramic.

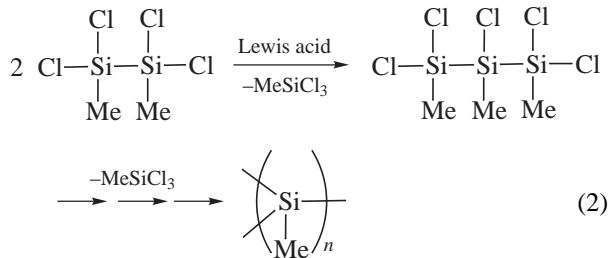
1. Organosilicon Polymers (Si–C–N–O Systems)

A wide variety of organosilicon preceramic polymers are accessible due to the satisfactory chemical stability of the covalent Si–C, Si–Si, Si–N, Si–O, and Si–H bonds.

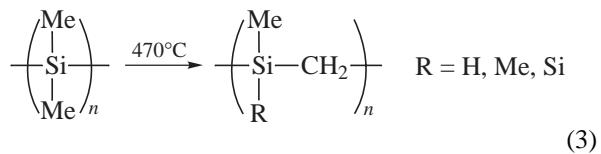
a. Polysilanes. Polysilanes have a backbone made of silicon atoms; they are usually prepared from inexpensive chlorosilane monomers by dehalogenation with sodium:



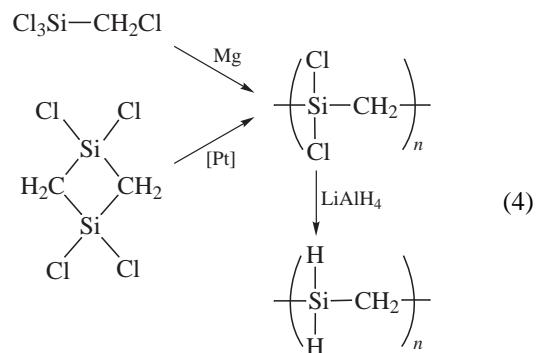
Polysilanes have also been prepared by redistribution of Si–Si and Si–Cl bonds of a mixture of chloromethyl-disilanes, by-products in the industrial synthesis of methylchlorosilanes. This redistribution leads to a branched polysilane with elimination of volatile trichloromethylsilane:



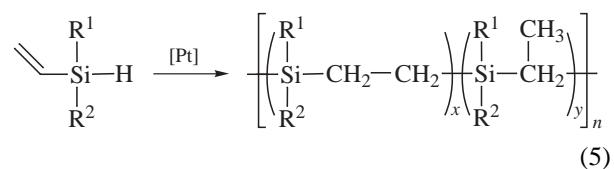
b. Polycarbosilanes. The polycarbosilane (PCS) used for the elaboration of the Nicalon™ SiC fiber, is prepared by heat treatment of polydimethylsilane (Kumada rearrangement) in an autoclave, which leads to the formation of a branched, polycyclic and/or cross-linked polycarbosilane:



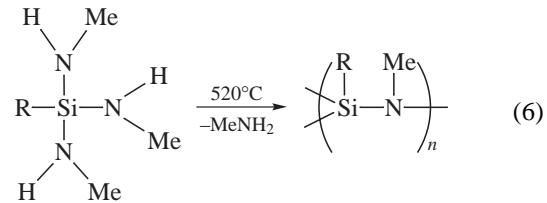
Other polycarbosilane precursors have been prepared using different syntheses. For instance, highly branched polysilaethylene ($-\text{SiH}_2-\text{CH}_2-$)_n can be prepared by coupling of chloromethyltrichlorosilane, and the linear polymer by ring-opening polymerization of disilacyclobutane monomers (Eq. (4)), in each case followed by reduction:



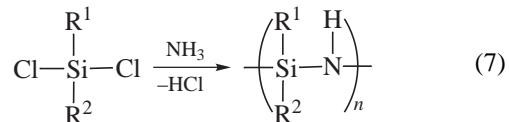
Polymerization of vinylhydrogenosilanes by hydrosilylation leads to polysilapropylene polymers with Si–CH₂CH₂–Si and Si–CH(CH₃)–Si carbosilane linkages resulting from β - and α -addition, respectively:



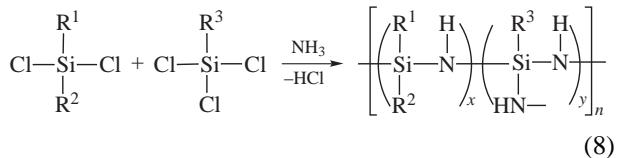
c. Polysilazanes and related polymers. The first polysilazane precursors to Si–C–N ceramics were produced by condensation (transamination) of aminosilanes:



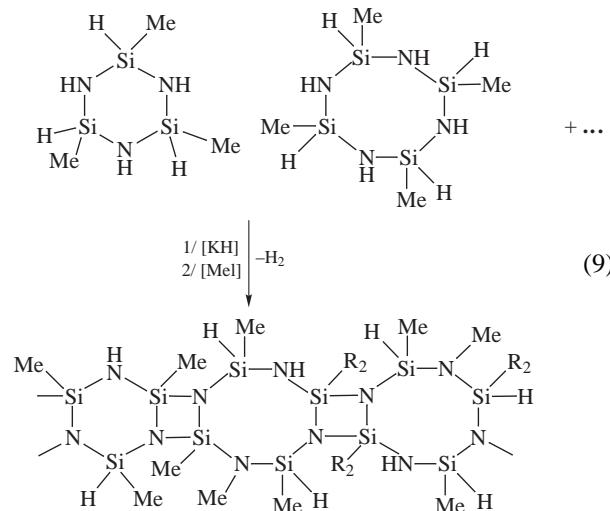
Polysilazanes are generally prepared by ammonolysis (or aminolysis) of chlorosilanes:



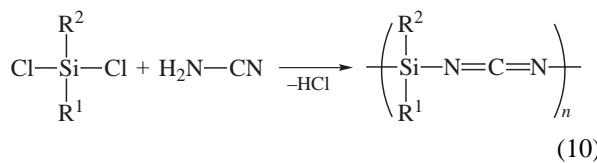
However, this reaction usually leads to low-molecular-weight linear and cyclic oligomers, too volatile to be used as precursors. Copolymerization with a trifunctional chlorosilane, leading to branched or cross-linked polymers, can be used to increase the molecular weight:



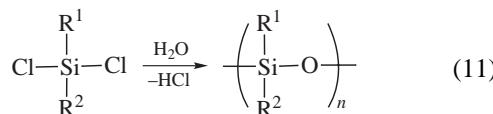
Alternatively, Si–H containing oligosilazanes can be cross-linked by dehydrocoupling between Si–H and N–H bonds. This reaction can be catalyzed by bases (Eq. (9)) or transition metal complexes such as Ru₃(CO)₁₂:



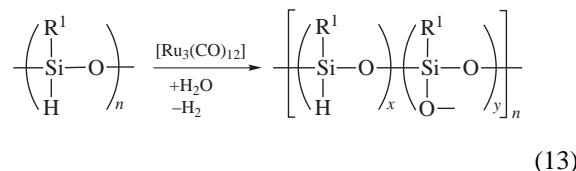
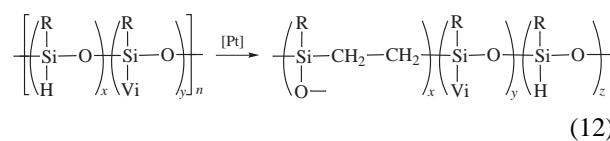
Polycarbosilazane precursors with Si, N, and C atoms in the backbone can be prepared by polycondensation of a chlorosilane monomer with a diamine $\text{H}_2\text{N}-\text{R}-\text{NH}_2$ monomer or by ammonolysis of chlorocarbosilane monomers $\text{Cl}\text{R}_2'\text{Si}-\text{R}-\text{SiR}'\text{Cl}$. More recently, polysilylcarbodiimide precursors have been synthesized by condensation of chlorosilanes with cyanamide:



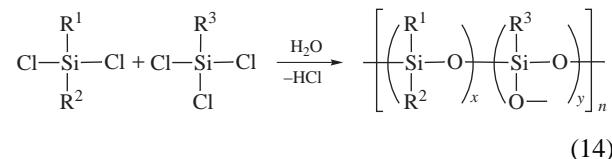
d. Polysiloxanes. Cross-linked polysiloxanes are inexpensive preceramic precursors for the preparation of silicon oxycarbide or oxynitride ceramics. Linear polysiloxanes are obtained by hydrolysis/condensation of dichlorosilanes (Eq. (11)), leading to low-molecular-weight cyclic oligomers that are converted into high-molecular-weight viscous polymers using acid or base catalysis:



To obtain suitable precursors, the oligomers and the polymers must be cross-linked. Hydrosilation of vinyl and Si-H containing polysiloxanes catalyzed by Pt compounds leads to polycarbosiloxane precursors (Eq. (12)). Dehydrocoupling of Si-H containing polysiloxanes with NH_3 or H_2O leads to polysiloxazane or polysiloxane precursors (Eq. (13)):



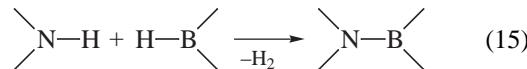
Alternatively, cross-linked polysiloxanes can be obtained directly using trifunctional or tetrafunctional monomers such as organotrichlorosilane (Eq. (14)) or tetrachlorosilane:



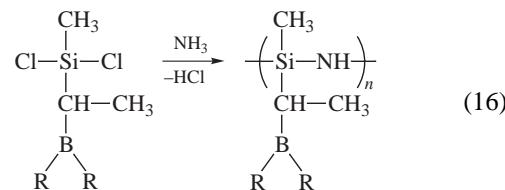
2. Introduction of Other Elements: The Si-C-E-N(O) System

The reactivity of the substituents attached to the silicon, nitrogen, or carbon atoms of the polymer backbone allows the introduction of another element, E (such as B, Al, or Ti) into the precursor and thus modification of the composition of the final ceramic. Alternatively, E containing organosilicon precursors can be prepared directly by polymerization of E-containing monomers.

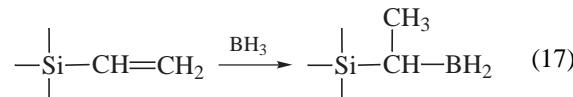
For instance, dehydrocoupling between N–H groups in polysilazanes and B–H groups in borane (BH_3) or borazine ($\text{B}_3\text{N}_3\text{H}_6$) leads to polyborosilazanes:



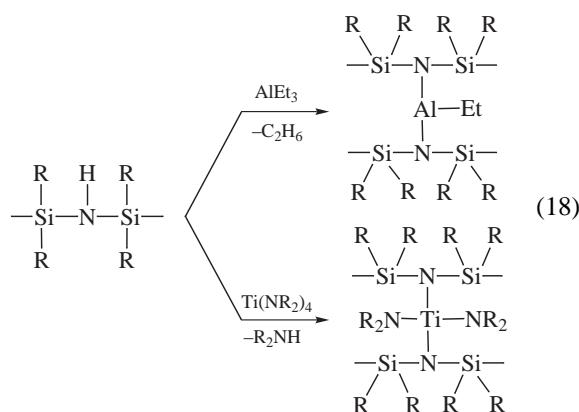
Alternatively, polyborosilazanes can be prepared by ammonolysis of a vinylchlorosilane monomer modified by hydroboration:



The hydroboration of vinyl substituents is also a convenient way to prepare boron containing precursors and ceramics:



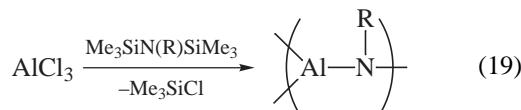
Polyaluminosilazanes and polytitanosilazanes have been prepared by reaction of N–H bonds in polysilazanes with trialkylaluminum and $\text{Ti}(\text{NMe}_2)_4$, respectively:



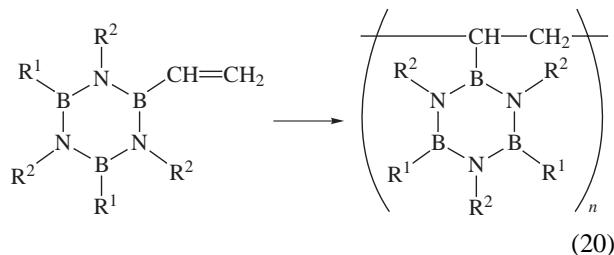
3. Silicon-Free Systems

Many efforts have been made to develop precursors to nonoxide, silicon-free ceramics such as BN, B–C, B–C–N, AlN, TiN, Ti–N–C, ZrN, and W₂C. The very high reactivity towards air or moisture of M–H and M–C bonds (M = Al, Ti, ...) makes it difficult to control the degree of cross-linking so as to prepare processable polymers, as well as making it difficult to handle both the monomers and the polymers.

Ammonolysis or aminolysis of chlorides MCl_n leads to cross-linked polymetalimides [M(NR)_{n/2}]_x that can be converted to the corresponding nitrides by pyrolysis in NH₃ (Table I). Redistribution between metal chlorides and hexamethyl disilazane provides an interesting route to M–N–(C) precursors:

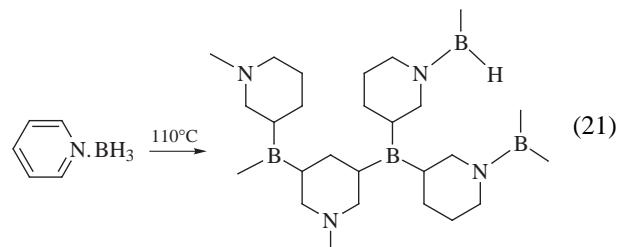


The situation is more favorable in the case of boron. Boranes and borazines provide attractive monomers. B–N–(C) precursors can be prepared by free-radical polymerization of vinylboranes or B–vinyl borazines, leading to vinylic polymers with pendent borane or borazine groups:

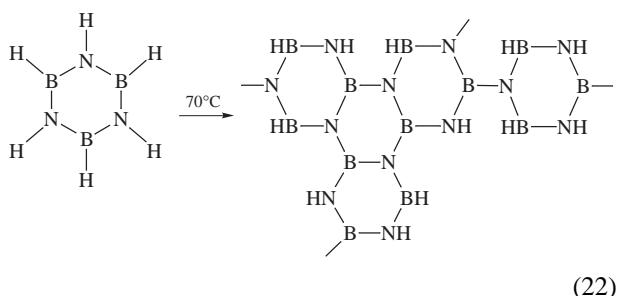


Another route to B–N–C precursors is the thermolysis of amino–borane complexes, such as piperazine–borane or ethylene diamine–borane complexes. In the particular

case of the pyridine–borane complex, thermal treatment at 110°C leads, via dehydrocoupling (Eq. (15)) and hydroboration (Eq. (17)) reactions, to a highly cross-linked precursor to B–N–C ceramics:

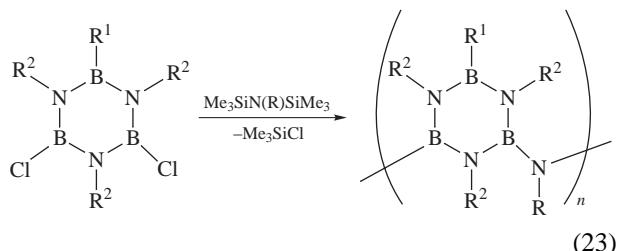


Thermal dehydrocoupling (Eq. (15)) between borazine B₃N₃H₆ monomers leads to a branched, polycyclic polyborazylene:



Other BN precursors have been obtained by thermal or metal-catalyzed dehydrocoupling (Eq. (15)) between ammonia or amines and various B–H-containing monomers.

B–N bonds can also be formed by condensation between N–H and B–Cl bonds or by redistribution between N–Si and B–Cl bonds. Thus, linear or cross-linked polyborazinylamines with controlled structures can be prepared by redistribution of di- or trichloroborazines with disilazanes:



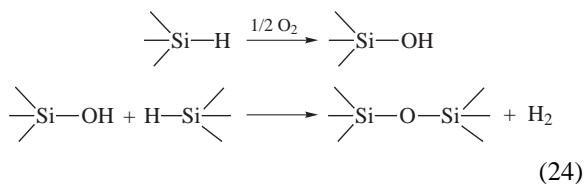
III. CROSS-LINKING AND PYROLYSIS

The polymeric precursors are converted into an inorganic ceramic by pyrolysis. The first concern in this step is the ceramic yield. Very few polymeric precursors give satisfactory ceramic yields if pyrolyzed directly after synthesis. This is due to cleavage of the polymer chains leading to the formation of volatile, low-molecular-weight compounds

that are lost during pyrolysis (see Section III.B). In addition, the precursors often contain a significant amount of short oligomers that are distilled during the first step of pyrolysis. The aim of the curing (or cross-linking) step is to connect the oligomer and polymer chains by strong, ionic-covalent bonds, thus transforming the precursor into a highly cross-linked network (Fig. 2). In such a network, more fragmentation reactions are needed to give volatile compounds, and the probability of combination of the fragments (the reverse reaction) is increased, which improves the ceramic yield. It is important to keep in mind that a very high degree of cross-linking is required to maximize the ceramic yield (e.g., one cross-link for one or two monomeric units). Additionally, the cross-linking renders the precursor unmeltable, which allows maintenance of the shape (e.g., fiber or monolith) during pyrolysis.

A. Cross-Linking

Curing can be achieved in principle by various means, such as oxidation in air, using chemical additives, irradiation, or thermally. For instance, curing of polycarbosilane-derived fibers (Nicalon, Tyranno) is performed by heating in air or in oxygen at 150 to 200°C (below the melting temperature of the fiber). This treatment results in the formation of very stable siloxane bridges between the chains, mostly by oxidation of the reactive Si–H bonds:



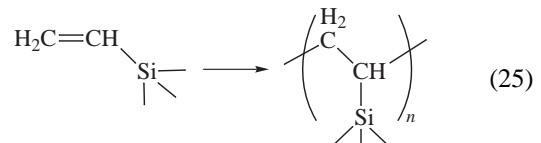
The main drawback of this procedure is that it introduces considerable amounts of oxygen (up to 14 wt%), which is not eliminated during subsequent pyrolysis. This high oxygen content limits the high-temperature applicability of these fibers to about 1200°C. To avoid the introduction of oxygen, other curing procedures have been explored, involving, for example, reaction with NO₂ and BCl₃, or CCl₄. Cross-linking by γ - or electron irradiation using electron accelerators or cobalt sources has also been employed. The effect of irradiation is likely to initiate the cleavage of Si–H and C–H bonds, the combination of the resulting free-radical intermediates leading to the formation of Si–Si, Si–C, or C–C cross-links.

One way to avoid the introduction of other elements or the use of complicated equipment is to prepare precursors with thermally reactive functionality, allowing the cross-linking of the polymer by a simple heat treatment. Thermal cross-linking appears very attractive, but it is somewhat difficult to control. The cross-linking temperature must be higher than the processing temperature but lower than

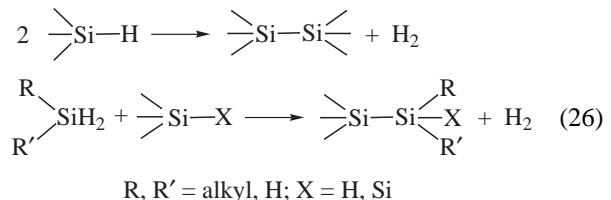
the degradation temperature of the precursor. In addition, if the aim of the curing is also to maintain the shape of fibers or monoliths, the cross-linking temperature must be lower than the melting temperature of the precursor, which obviously prevents any melt processing. Thus, the control of the cross-linking temperature by catalysis of the cross-linking reactions may be necessary.

In the case of organosilicon polymers, vinyl, Si–H, and N–H (in polysilazanes) groups have been used extensively. Depending on the reactive groups present, several cross-linking reactions are possible:

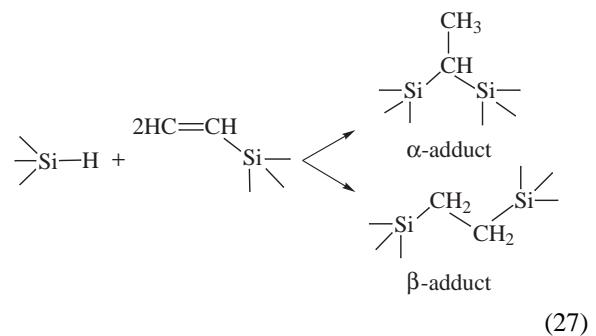
- Vinyl polymerization (may be catalyzed by peroxides, ultraviolet irradiation, etc.):



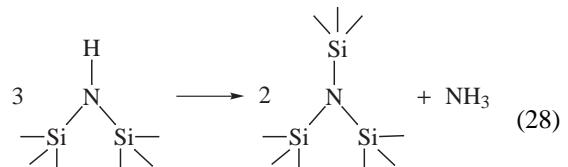
- Dehydrogenation involving Si–H or SiH₂ groups (Eq. (26), may be catalyzed by Ti or other transition metal compounds):



- Hydrosilylation (may be catalyzed by Pt or other compounds):



- Transamination:



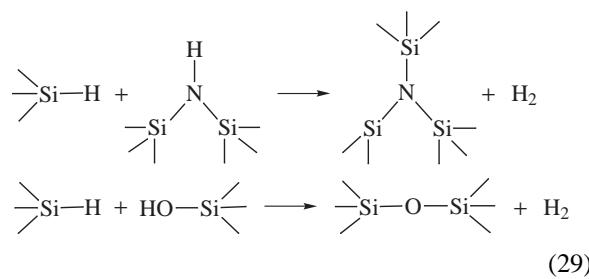
- Dehydrogenation (dehydrocoupling) between Si–H and Si–X–H groups, X = N, O (may be catalyzed by KH or Ru3(CO)12):

TABLE II Loss of Silicon During Pyrolysis and Ceramic Yields for Silazane Oligomers and Polycarbosilanes with Different Functionalities

Precursor (theoretical formula)	Loss of Si (%)	Ceramic yield (%)
$(-\text{SiViH}-\text{NH}-)_n^a$	0.5	83
$(-\text{SiViH}-\text{NMe}-)_n^a$	13	66
$(-\text{SiMeH}-\text{NH}-)_n^a$	46	46
$(-\text{SiViH}-\text{NH}-)_x (-\text{SiMe}_2-\text{NH}-)_y^a$	17	63
$(-\text{SiMe}_2-\text{CH}_2\text{CH}_2-)_n^b$	100	0
$(-\text{SiH}_2-\text{CH}_2\text{CH}_2-)_n^b$	48	38
$(-\text{SiH}_2-\text{CH}_2-)_n^b$	8	85

^a Cured at 120°C.

^b No curing.



The vinyl groups are particularly efficient for thermal cross-linking, but they increase the amount of excess carbon in the ceramic. Conversely, the use of Si–H groups permits one to lower the excess carbon content. The thermal reactivity of SiH₂ groups is much higher than that of Si–H groups, due to the facile formation of silylene (=Si•) rather than free-radical (=Si[•]) intermediates.

Comparison of the silicon losses during the pyrolysis and ceramic yields for silazane oligomers and polycarbosilanes with different functionalities illustrates very well the importance of the functionality of the precursor (Table II).

B. Pyrolysis

The pyrolysis permits conversion of the organosilicon or organometallic precursor into an inorganic ceramic. This step involves many complicated reactions, compositional modifications, and structural rearrangements. The main parameter is the nature of the pyrolysis atmosphere, inert or reactive, which permits one to change to a considerable extent the final ceramic composition. The temperature program used for the pyrolysis is also important. For instance, not only the crystallinity but also the composition of the ceramic may depend on the final temperature and annealing time, and the ceramic yield may depend on the heating rate.

An understanding of the pyrolysis of polymeric precursors requires the use of many complementary techniques, such as:

- Thermogravimetric analysis coupled with mass spectrometry (TG/MS) and gas chromatography coupled with mass spectrometry (GC/MS) for the analysis of the volatile products

- elemental analysis; spectroscopies such as infrared (IR), nuclear magnetic resonance (NMR); X-ray photoelectron spectroscopy (XPS); X-ray diffraction (XRD); and transmission electron microscopy (TEM) for the analysis of the ceramic residue

In the following section, we will focus on the pyrolysis chemistry of organosilicon precursors, which has been extensively studied.

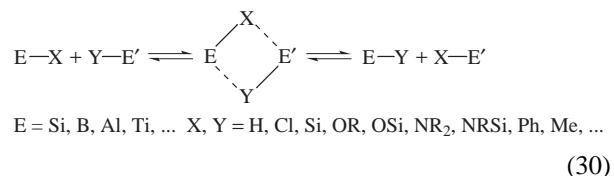
1. Inert Atmosphere

Four stages may be distinguished in the pyrolysis under argon of an organosilicon precursor between room temperature and 1500°C (Fig. 3):

- Distillation of volatile oligomers (if any)
- Degradation, often leading to the loss of organosilicon compounds
- Mineralization, with loss of hydrocarbons and dihydrogen
- Decomposition (Si–O- or Si–N-containing ceramics), with loss of carbon monoxide, silicon monoxide, or dinitrogen.

a. Degradation stage. Between about 300 and 600°C, organosilicon precursors undergo degradation reactions involving cleavage of the polymeric backbone. In the same temperature range, thermal cross-linking reactions involving the various functional groups on the precursor also take place (see Section III.A). The degradation reactions depend on the nature of the the precursor.

Redistribution reactions (also called exchange reactions) are very common reactions in organometallic chemistry; they lead to the exchange of the substituents (H, Cl, OR, etc.) around a central atom (Si, B, Al, etc.):



Redistributions are catalyzed by acids or bases or are thermally activated. It is not surprising, then, that they play a prominent part in the pyrolysis of all organoelement precursors. Thus, redistributions are involved in the degradation of polysilanes, polysiloxanes, and polysilazanes.

The redistribution of polysiloxane precursors illustrates very well the importance of these reactions. The

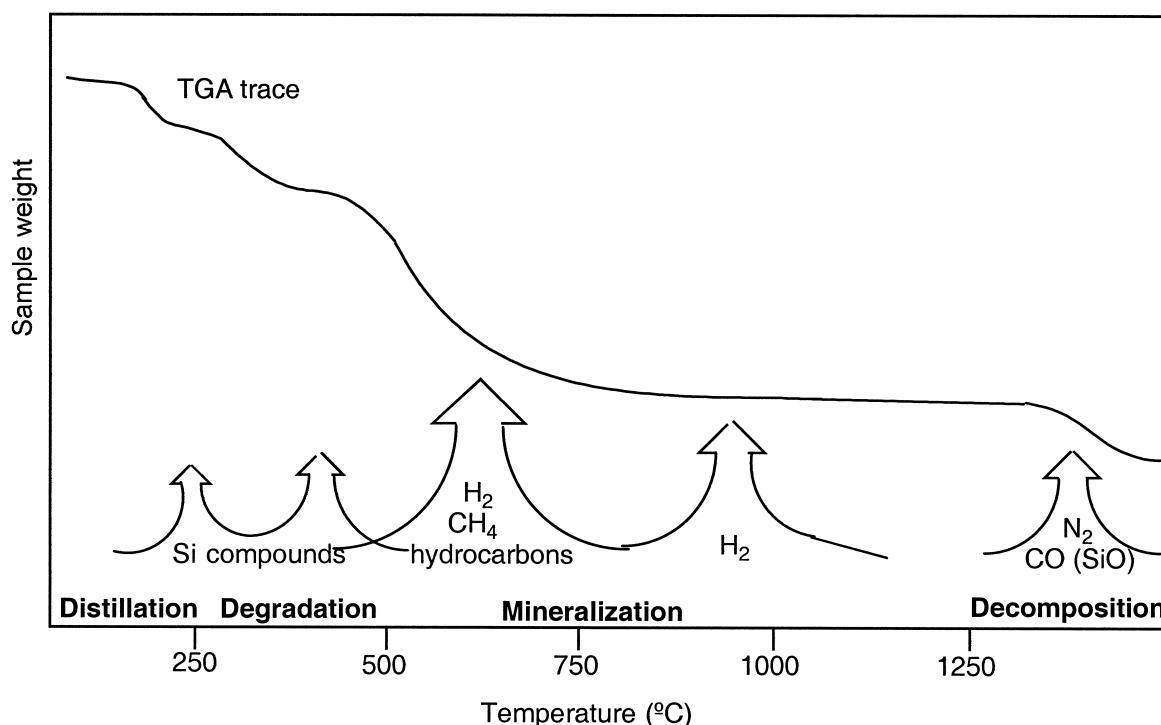
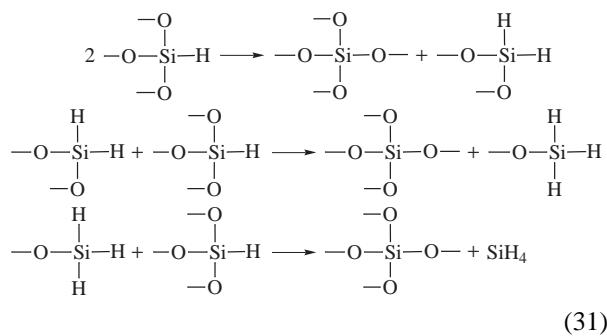


FIGURE 3 Typical thermogravimetric analysis and gas evolution during the pyrolysis of an organosilicon precursor (inert atmosphere).

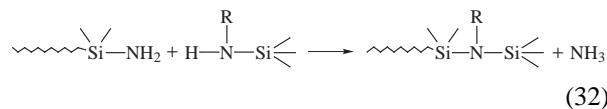
redistribution of Si–O and Si–X bonds (X = H, Si≡, alkyl, aryl, O–Si≡) during pyrolysis of cross-linked polysiloxanes has many implications. Successive redistribution steps lead to the escape of volatile organosilicon compounds such as silanes or linear and cyclic oligosiloxanes. The loss of these compounds decreases the ceramic yield and increases both the O/Si content and the cross-linking in the residue. For instance, in $HSiO_{1.5}$ precursors, a significant amount of silane SiH_4 is formed by Si–O/Si–H redistribution:



Even if the loss of organosilicon compounds is negligible, the redistribution reactions completely modify the environment of the Si atoms, leading in the final Si–O–C (inert atmosphere) or Si–O–N (ammonia atmosphere) ce-

ramics to a distribution of SiO_xC_{4-x} or SiN_xC_{4-x} sites (Fig. 4).

Nucleophilic attack by OH or NH groups are also involved in the degradation processes of polysiloxanes and polysilazanes. The latter reaction, also called transamination, implies the cleavage of Si–N bonds and may lead to the escape of volatile aminosilanes or oligosilazanes. On the other hand, successive transamination steps lead to the loss of NH_3 , which lowers the final N/Si ratio and increases the degree of cross-linking:



Si–C/Si–C or Si–C/Si–H redistributions have high activation energies. As a result, polycarbosilanes are more stable than other organosilicon precursors, and the degradation takes place at the beginning of the mineralization stage (see Section III.B.1.b). The main degradation pathway likely involves the homolytic cleavage of Si–C bonds, followed by hydrogen abstraction or rearrangements. For instance, in the case of PCS, the precursor of the Nicalon fiber, successive cleavage of the Si–CH₂–Si bridges and hydrogen abstraction lead to the escape of a mixture of methylsilanes Me_xSiH_{4-x} :

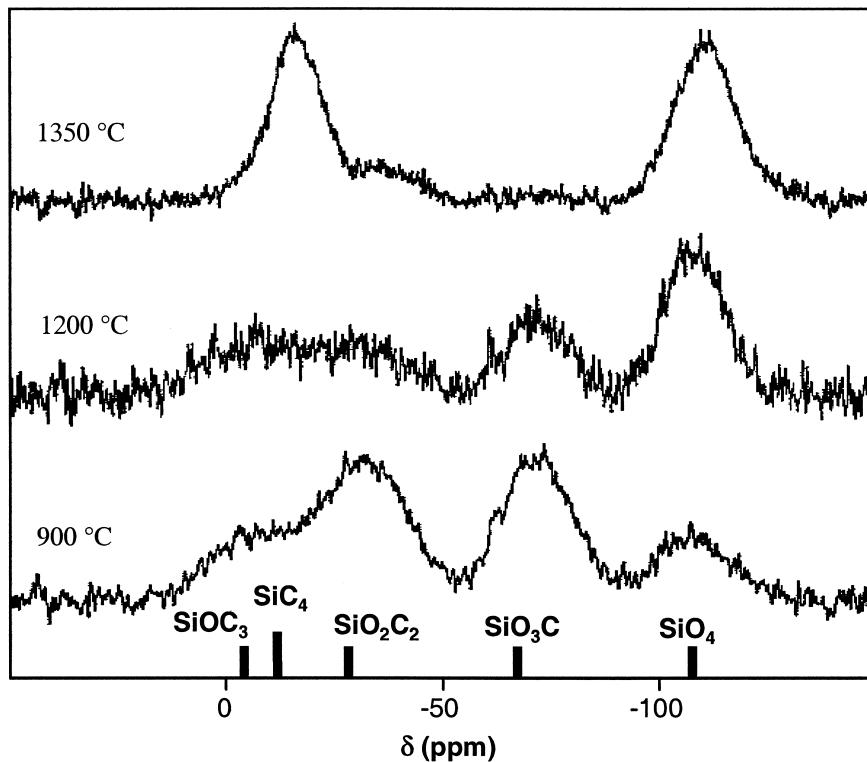
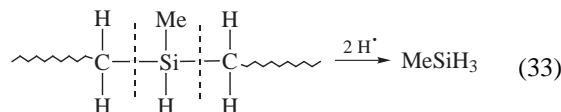
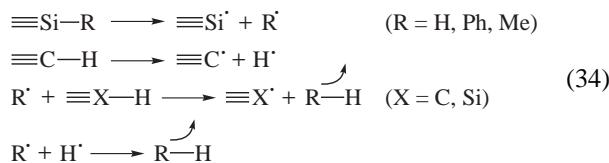


FIGURE 4 Solid-state ^{29}Si NMR spectrum of a Si–O–C ceramic (composition: $\text{SiO}_{1.14}\text{C}_{1.74}$) derived from a cross-linked polysiloxane precursor pyrolyzed in argon at 900°C (30 min), 1200°C (1 hr), and 1350°C (24 hr).



Simultaneously, the formation of Si–C and C–C bonds by combination of the free radicals increases the degree of cross-linking in the residue.

b. Mineralization stage. Between about 400 and 1000°C , extensive cleavage of N–C, N–H, Si–C, C–C, and then C–H bonds takes place, leading to the escape of hydrocarbons and hydrogen and to the formation of an inorganic, amorphous residue. The formation of these gases may be accounted for by homolytic bond cleavage, followed by hydrogen abstraction, rearrangement, or combination of the free-radical intermediates, as exemplified in Eq. (34):

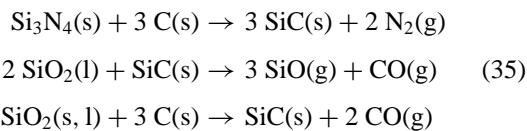


The escape of these gases leads to a decrease in the C/Si and H/Si ratio of the residue. As the temperature increases,

the rate of bond cleavage increases and hydrogen atoms become less and less abundant. Si–H and C–H bonds gradually disappear, replaced by Si–C bonds formed by combination between the Si^{\cdot} and C^{\cdot} free radicals. Simultaneously, the formation of C–C bonds leads to a phase of polyaromatic carbon (free carbon) similar to the carbon obtained by pyrolysis of organic polymers. The amount of free carbon depends on the nature of the organic groups: methyl groups or linear alkyl groups lead to moderate free-carbon contents, while unsaturated groups lead to higher free-carbon contents.

At the end of this stage, disordered amorphous covalent ceramics are obtained. These amorphous ceramics are essentially built of the same Si–C, Si–O, or Si–N bonds as the stable SiC , SiO_2 , or Si_3N_4 phases, but a small percentage of C–H, O–H, or N–H bonds are still present. In binary systems (Si–C or Si–N), lack of long-range order arises from the presence of H atoms and from slight variations of bond lengths and bond angles, as in silica glass. In ternary (or multinary) systems, Si–O–C, Si–O–N, or Si–C–N, the disorder concerns also the environment of the Si atoms, which may be bonded to different atoms in mixed $\text{SiC}_x\text{X}_{4-x}$ tetrahedra ($\text{X} = \text{O, N}$) (Fig. 4). This mixed environment results from the occurrence of redistribution reactions in the 400 to 1000°C range.

c. Decomposition stage. The ceramics obtained at 1000°C after the mineralization stage are not necessarily stable at higher temperatures. Thus, above about 1400°C an additional weight loss is observed for Si–O- or Si–N-containing precursors. This weight loss corresponds to the escape of CO, Si–O, or N₂, usually arising from reactions between the different phases present in the ceramic, such as those given in Eq. (35):



The latter reaction, also called carbothermal reduction, may be used to prepare silicon carbide from polysiloxane-derived Si–O–C ceramics, according to the overall equation:



Tuning the C content (C/Si = 1 + O/Si) of the intermediate Si–O–C ceramic is achieved by the use of vinyl or phenyl substituents in the precursor.

All of these reactions limit the high-temperature stability of the ceramics. The incorporation of boron considerably enhances the high-temperature stability of Si–C–N

ceramics. This may be ascribed to a decrease of the atomic mobility in Si–B–C–N ceramics and to inhibition of grain growth by formation of BN layers at the grain boundaries of the crystals. Another way of improving the high-temperature stability is to precisely control the composition of the ceramic. For instance, if the ceramic does not contain free carbon or oxygen, the decomposition reactions given in Eq. (35) are no longer possible.

2. Ammonia Atmosphere

The pyrolysis in ammonia of an organosilicon precursor completely modifies the composition of the ceramic. Below about 750°C, substitution by NH₃ converts the Si–H, Si–Si, and Si–C bonds into Si–N bonds (nitridation stage), while the Si–O bonds are unaffected. If free carbon forms, it is removed above 750°C as hydrogen cyanide (free-carbon removal stage; see Fig. 5). Distillation and degradation stages are similar to those observed in an inert atmosphere, although the organosilicon compounds evolved may contain NH₂ substituents, indicating partial nitridation in the same temperature range.

a. Nitridation and free-carbon removal stages. In this stage, the Si–C, Si–H, and Si–Si bonds are replaced

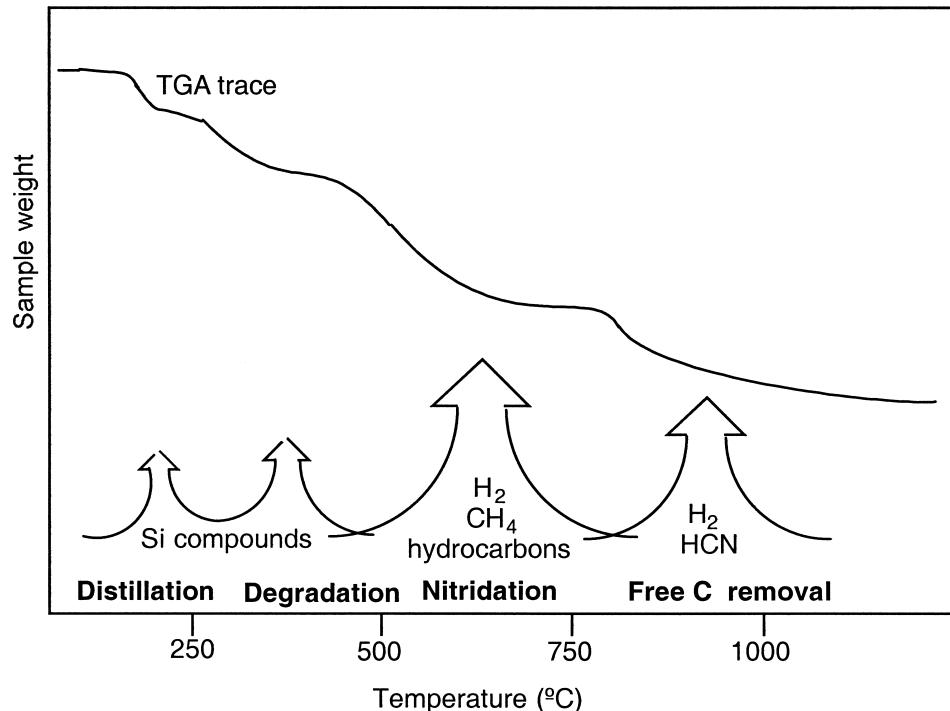
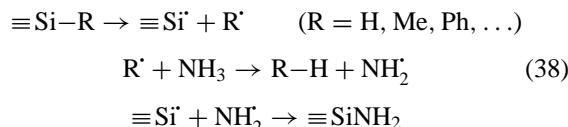
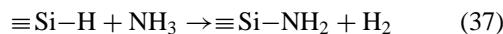


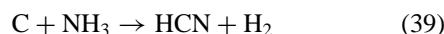
FIGURE 5 Typical thermogravimetric analysis and gas evolution during the pyrolysis of an organosilicon precursor (NH₃ atmosphere).

by Si–N bonds. The Si–O bonds are not modified. The formation of Si–N bonds results from nucleophilic (Eq. (37)) and free-radical substitution by ammonia (Eq. (38)), leading to Si–NH₂ groups as intermediates:



Condensation of these Si–NH₂ groups (transamination, Eq. (32)) leads to Si₂NH and finally to Si₃N groups.

Ammonia provides an infinite source of hydrogen, and the probability of H abstraction by the R· free radicals is very high, which hinders the formation of free carbon. For instance, methyl groups are quantitatively converted to methane. If free-carbon forms (precursors with vinyl or phenyl substituents, or containing Si–CH₂–CH₂–Si bridges), it is removed above 750°C as hydrogen cyanide (free-carbon removal stage):



IV. CRYSTALLIZATION

Above 1000°C, the disordered amorphous ceramics crystallize into the thermodynamically stable crystalline phases. In silicon-based ceramics, crystallization leads to a complete modification of the environment of the Si atoms: the Si atoms with a mixed environment (e.g., SiC_xX_{4-x}, with 1 ≤ x ≤ 3, X = O, N) gradually rearrange to the SiC₄, SiO₄, and Si₃N₄ tetrahedra found in the stable SiO₂, SiC, and Si₃N₄ phases (Fig. 4).

The crystallization behavior depends strongly on the annealing temperature and on the composition of the amorphous ceramic and also on the relative stability of the crystalline phases (see Sections III.1.c). Diffusion and grain growth lead to nano/nano, nano/micro, or micro/micro composites. The crystallite size and crystalline content increase as the pyrolysis temperature increases and as the bulk composition approaches the stoichiometry of the stable phases (e.g., SiC or Si₃N₄). At intermediate temperatures, mixtures of amorphous and crystalline phases are obtained, leading to a complicated nanostructure. Thus, the Nicalon Si–O–C fibers consist of a continuous silicon oxycarbide phase, containing dispersed carbon and β-SiC nanocrystallites as well as closed nanopores. As mentioned in Sections III.1.c and V.1, the incorporation of boron significantly modifies the crystallization behavior of silicon-based ceramics. Doping of the amorphous ceramic

with titanium also retards the crystallization, whereas doping with phosphorus accelerates it.

V. PRECURSOR PROCESSING AND APPLICATIONS

One of the main advantages of the precursor pyrolysis route is the possibility of achieving a variety of ceramic shapes: fibers, monoliths, films, or powders. Shaping is usually performed at the precursor stage; the precursor is then transformed into an unmeltable solid by cross-linking (curing), which allows maintenance of the shape during the final pyrolysis step.

Chemical synthesis (choice of the monomers, polymerization step) provides a wide variety of precursors with different chemical and physical properties. The choice of the shaping technique mainly depends on the meltability, the solubility, or the volatility of the precursor (Fig. 6). Besides these properties, other properties are important for the processing:

- Rheology (viscosity)
- Thermal stability (melting without decomposition)
- Chemical stability (stable viscosity, no gel formation on aging, inertness towards air or atmospheric moisture)

1. Fibers

Fine-diameter ceramic fibers suitable for weaving and knitting can be obtained by melt- or dry-spinning of meltable or soluble precursors. Besides an appropriate rheology, a high thermal stability and a high purity (absence of particulates) are needed to obtain high-quality fibers. High functionality and reactivity of the precursor allowing rapid and extensive cross-linking are essential. Cross-linking are required to make the fibers unmeltable so they can survive the pyrolysis step and to maximize the ceramic yield.

Most fibers reported are based on silicon carbide or silicon nitride (Si–C–O, Si–N, Si–C–N–O). Large amounts of oxygen (due to the cross-linking process) and excess carbon are often present. Thus, the well-known Nicalon fiber, obtained by pyrolysis of a polycarbosilane contains about 14 wt% oxygen and 12 wt% excess carbon. Both oxygen and free carbon decrease the operating temperature of the fiber to 1200°C. This temperature is increased to around 1300°C by using a polytitanocarbosilane precursor (Tyranno Si–Ti–C–O fibers). Boron-containing fibers (Si–C–B, Si–B–C–N, Si–B–C–N–O), obtained by treatment of the precursor fiber with BCl₃ prior to pyrolysis or from organoborosilicon polymers, show excellent

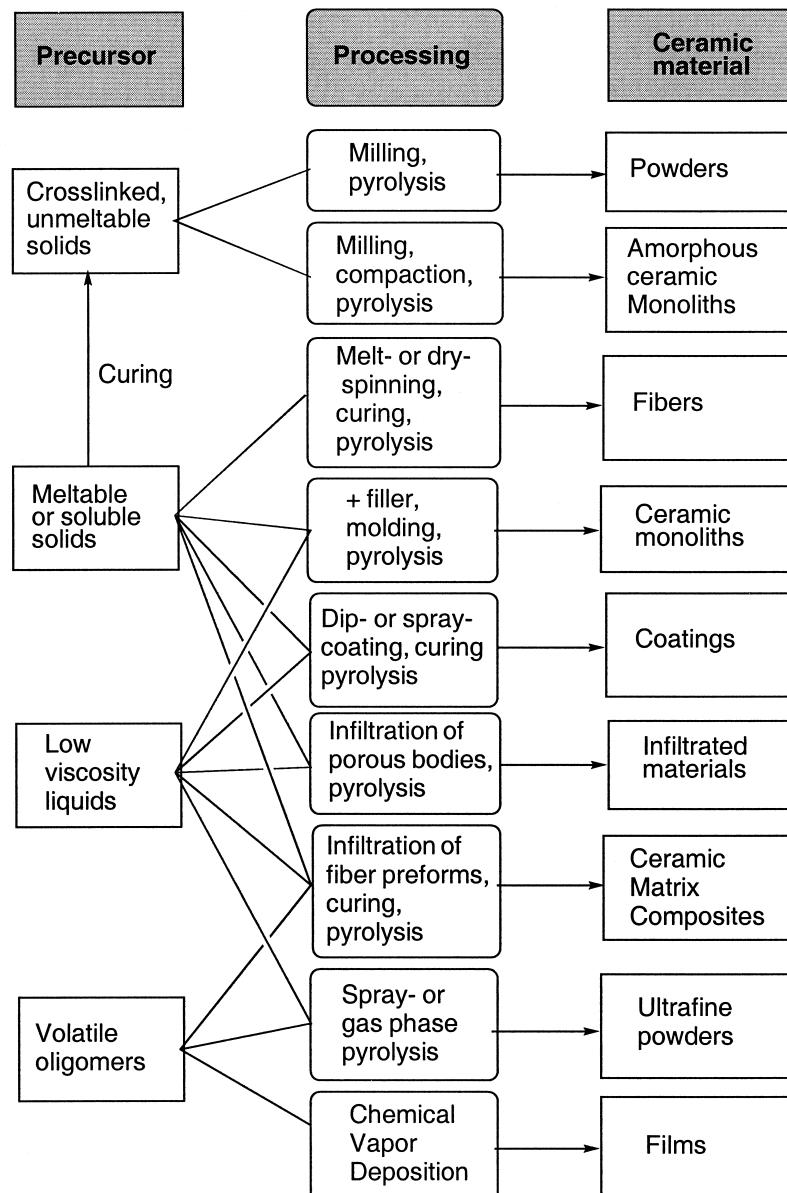


FIGURE 6 Processing and applications of polymeric precursors.

thermomechanical properties and oxidation stability, up to about 1600°C. The properties of the reported ceramic fibers are in the following ranges:

- Diameter 8 to 20 μm
- Density 1.8 to 3.1 g/cm³
- Tensile strength 2 to 4 GPa
- Elastic modulus 200 to 430 GPa

2. Binders and Sintering Aids

Preceramic polymers are promising forming aids for the fabrication of dense ceramic components from con-

ventional nonoxide ceramic powders (fillers). Moldable precursors (meltable or soluble) are attractive binders for the injection molding of ceramics. The polymer acts as a binder and a sintering additive. One advantage of using a binder system that can be converted to ceramic material is the resulting increase in green density upon pyrolysis of the polymer, especially when the powder has poor packing characteristics. This leads to less shrinkage and enhanced mechanical properties. Instead of passive ceramic powders, metal powders (Al, B, Si, MoSi₂, etc.) can be used. These powders (active fillers) can react with the precursor, its decomposition products, or the atmosphere during pyrolysis, leading to carbide, nitride, or oxide products

which reduces the loss of volatile products and the shrinkage.

3. Infiltration of Porous Bodies

The infiltration of porous matrices (e.g., carbon, reaction-bonded silicon nitride) by a soluble or liquid precursor (with a high ceramic yield) followed by pyrolysis results in increased densities and strength and decreased open porosity, thus improving oxidation resistance.

4. Coatings

Preceramic polymers can be used for the formation of protective coatings on other ceramics, metals, or carbon. The simplest process is dip-coating, which consists of dipping the substrate into the precursor solution or melt and withdrawing it at a controlled rate. In this process, the film thickness depends on the withdrawal rate, the viscosity of the melt, or the concentration of the polymer solution. The polymer coating can then be transformed by pyrolysis into strongly adhering, amorphous ceramic coating; however, volume-shrinkage during the pyrolysis leads to crack formation above a certain critical layer thickness. To obtain thick layers, one can use multiple dip-coating/pyrolysis cycles or add passive or active fillers (e.g., ceramic or metal powders) to the preceramic polymers, prior to the dip-coating process, in order to minimize shrinkage during pyrolysis. These coatings serve to improve the hardness and the oxidation resistance of the substrate and can act as electrical and thermal insulation or as a binder phase.

5. Amorphous Covalent Ceramic Monoliths

The possibility of transforming precursor green bodies (bodies made of compacted precursor powder) directly into ceramic monoliths is an extremely promising and elegant way to avoid the use of high-temperature, high-pressure, and sintering aids necessary in conventional densification of ceramic powders. In this process, the precursor must be a highly cross-linked unmeltable polymer to avoid bloating during the pyrolysis and to maximize the ceramic yield; however, it must still contain enough functionality to densify during compaction and pyrolysis. After milling and sieving, the precursor powder is warm-pressed to a dense green body and finally pyrolyzed at around 1000°C, to yield an amorphous covalent ceramic (ACC) monolith. Warm-pressing above the glass-transition temperature of the precursor is important to maximize the density of the final ACC. Thus, thermally cross-linked polysilazane precursors have been converted to crack-free amorphous Si-C-N materials with relative density >97%

and superior high-temperature stability. These amorphous ceramics are homogeneous on the atomic scale; on annealing above 1400°C, phase separation of the crystalline stable phases leads to $\text{Si}_3\text{N}_4/\text{SiC}$ micro- or nanocomposites.

6. Matrices

Ceramic matrix composites (CMCs) may be obtained by liquid- or gas-phase infiltration of carbon or ceramic fiber preforms with a precursor, followed by thermal cross-linking in an autoclave and pyrolysis. The porosity of the resulting material may be reduced by a further infiltration, cross-linking, and pyrolysis step. CMCs can be used to overcome the brittle, catastrophic failure of conventional monolithic ceramic materials and are among the most promising thermostructural materials for ultra-high-temperature applications (above 1300°C), in the aerospace industry as well as the automotive and energy technologies.

7. Powders

Submicron ceramic powders have been obtained by gas-phase pyrolysis of volatile compounds and by spray pyrolysis of liquid precursors. Spray pyrolysis involves the formation of an aerosol by ultrasonication and pyrolysis using hot-tube, plasma, or laser heating. Smooth, spherical Si_3N_4 particles with a diameter of 38 nm have been obtained by this technique.

Amorphous ceramic powders derived from precursor pyrolysis can be used for the fabrication of ceramic micro- or nanocomposites. Densification of the powders is achieved by liquid-phase sintering, using sintering aids (Al or Y oxides). These sintering aids can be incorporated during the synthesis of the precursor using alkoxides of Al or Y. Complicated reactions involving the various phases in contact during the sintering lead to ceramic micro- or nanocomposites with a variety of microstructures and exceptional mechanical properties.

VI. OUTLOOK

At this time, the ceramics prepared by the precursor pyrolysis process are still expensive, not only because the preceramic polymers themselves are expensive but also because they are frequently air sensitive, which complicates their handling and processing and can dramatically increase the fabrication costs. Despite their unique shapes and microstructures, the use of precursor-derived ceramics has been mostly limited to high-tech applications, where their specific properties counterbalance their price, or to applications where the quantity of precursor required is

relatively small, as in coating, joining, etc. However, in some cases, such as the preparation of ceramic matrix composites, the price of precursors is coming down to the point where they are competing effectively with chemical vapor impregnation and other methods.

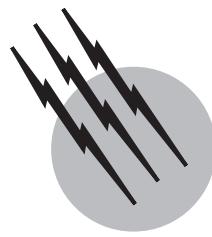
SEE ALSO THE FOLLOWING ARTICLES

CHEMICAL PROCESS DESIGN, SIMULATION, OPTIMIZATION, AND CONTROL • COMPOSITE MATERIALS • GLASS-CERAMICS • MANUFACTURING PROCESSES TECHNOLOGY • PHOTOCROMIC GLASSES • POLYMER PROCESSING

BIBLIOGRAPHY

Bill, J., and Aldinger, F. (1995). "Precursor-Derived Covalent Ceramics," *Advanced Materials* **7**, 775–787.

- Birot, M., Pillot, J.-P., and Dunogues, J. (1995). "Comprehensive chemistry of polycarbosilanes, polysilazanes, and polycarbosilazanes as precursors of ceramics," *Chemical Reviews* **95**, 1443–1477.
- Kroke, E., Li, Y.-L., Konetschny, C., Lecomte, E., Fasel, E., and Riedel, A. (2000). "Silazane-derived ceramics and related materials," *Materials Science & Engineering* **R26**, 95–199.
- Lipowitz, J. (1991). "Structure and properties of ceramic fibers prepared from organosilicon polymers," *Inorganic and Organometallic Polymers* **1**, 277–297.
- Mutin, P. H. (1999). "Control of the composition and structure of silicon oxycarbide and oxynitride glasses derived from polysiloxane precursors," *Journal of Sol-Gel Science and Technology* **14**, 27–38.
- Paine, R. T., and Narula, C. K. (1990). "Synthetic routes to boron nitride," *Chemical Reviews* **90**, 73–91.
- Paine, R. T., and Sneddon, L. G. (1994). "Borazine-based polymers close in on commercial performance," *CHEMTECH* 29–37.
- Peuckert, M., Vaahs, T., and Brück, M. (1990). "Ceramics from organometallic polymers," *Advanced Materials* **2**, 398–404.
- Seyferth, D. (1995). "Preceramic polymers: past, present and future." In "Materials Chemistry—An Emerging Discipline" (Interrante, L. V., Caspar, L. A., and Ellis, E. B., Editors), pp. 131–160, American Chemical Society, Washington, D.C.



Coatings, Colorants, and Paints

Peter K. T. Oldring

The Valspar Corporation

- I. History of Surface Coatings
- II. Natural Resins
- III. Oils
- IV. Ester Gums
- V. Oleoresinous Varnishes
- VI. Alkyds
- VII. Oil-Free Polyesters
- VIII. Acrylic Resins and Styrene
- IX. Phenoplasts
- X. Aminoplasts
- XI. Epoxy Resins
- XII. Polyurethanes
- XIII. Vinyl Acetate Polymers
- XIV. Vinyl Chloride and Vinylidene Chloride Polymers

- XV. Cellulose Derivatives
- XVI. Rubber and Its Derivatives
- XVII. Silicones
- XVIII. Polyolefins
- XIX. Pigments and Colorants
- XX. Solvents
- XXI. Radiation Curable Coatings
- XXII. High Solids Coatings
- XXIII. Powder Coatings
- XXIV. Waterborne Coatings
- XXV. Toxicity
- XXVI. Future Trends

GLOSSARY

This glossary is for paints and coatings and is limited to paints and coatings.

Alkyd Polyester resin containing vegetable oil (or derived fatty acids) and phthalic anhydride. Alkyds can be air drying or non-air drying.

Binder Film-forming material which “binds” the pigment in the coating composition and to the substrate.

Binders are part of a paint formulation and are in essence the non-volatile portion of the medium, colorants, additives, and fillers excluded.

Coalescence Fusing of soft, resin particles of an emulsion to form a continuous film.

Coating Pigmented or unpigmented film containing organic materials applied to a substrate such as wood, metal, plastic, paper, ceramic, or textile for decoration or protection. The term “coating” also includes paints, enamels, varnishes, lacquers, sizes, primers, basecoats, topcoats, tiecoats, and undercoats. It excludes adhesives, inks, and preformed films.

Colorant Pigment or dye which adds color to an ink or coating.

Crosslinking Chemical reactions which form a three-dimensional network in the binder, thereby giving the coating film its properties. Applicable only to thermoset coatings.

Crosslinking agent A chemically reactive species, often

a low molecular weight substance (often a resin) or a catalyst added to a thermoset resin in order to achieve crosslinking of the thermoset resin.

Cure The physical state of the film when it has achieved its desirable properties. Only applicable to thermoset coatings.

Dispersion Organic (e.g., polymer) or inorganic (e.g., pigments) particles dispersed in a medium, either aqueous or non-aqueous.

Drier Organometallic salt(s) added to air drying coatings which contain oils (oleoresinous or alkyds) to induce and promote the drying of the film through catalytic oxidation of fatty acid unsaturation of an alkyd or oleoresinous material. Different metal salts may have different functions, such as promoting surface or through drying.

Dye Colorant, which is soluble in the solvents and resin system, normally used in inks or textiles to produce transparent colored films.

Emulsion Suspension of fine particles of a liquid in a liquid, such as oil in water. Emulsion is frequently used to refer to polymeric dispersions in water, often prepared by emulsion polymerization.

Enamel Normally refers to a high-gloss pigmented coating on metal, glass, or ceramics. Enamels are coatings.

Film Dry or wet and not preformed. A liquid coating forms a wet film on the surface of the substrate when applied. When it becomes a coating, a continuous layer, (dry film) covering the required parts of the substrate, is formed.

Filler Insoluble particles in a coating which do not add color, but are added to improve the barrier effect of the coating film.

Gum Dried exudate of plants.

Ink Pigment dispersions in resins to provide external decoration. They are applied by printing processes directly onto substrates or onto coated substrate. Inks can be overvarnished.

Lacquer The term lacquer has different meanings in different industries and in different parts of the world. Originally, lacquers were based upon solutions of natural resins, which may or may not have been pigmented. Nowadays, lacquers are predominately transparent coatings based upon synthetic resins. Lacquers are coatings, being a continuous film formed by evaporation of solvent from a solution or dispersion.

Latex/Latices Originally referred to a natural rubber dispersion. Today, it is also used for synthetic polymer dispersions and emulsions, such as poly(vinyl acetate).

Lining Lining normally refers to can coatings and is the internal coating or lacquer (i.e., the coating in contact with foodstuffs). Lining is also used for interior tank coatings.

Medium The liquid phase of a paint.

Oil Fatty acid triglyceride, normally of vegetable origin. Oils are classified as drying, non-drying, and semi-drying.

Oil, drying Drying oils (normally vegetable oils) contain sufficient fatty acid unsaturation for free radical polymerization in the presence of air and a dryer to take place. An example is linseed oil.

Oil, non-drying Non-drying oils contain fatty acids with little or no unsaturation. An example is coconut oil.

Oil, semi-drying Semi-drying oils are intermediate to drying and non-drying. An example is soya bean oil.

Oleoresinous Containing a drying oil, usually in conjunction with a natural or synthetic resin.

Paint Mixture of pigment(s), resin(s), additives, and solvent(s) which form film when dry. Paint is a coating.

Pigment Synthetic or natural colorant, which is insoluble in the coating resins and solvents. A pigment needs dispersing and normally produces opaque-colored films. Other pigments include anticorrosive pigments, zinc dust, and aluminium flakes.

Polymer dispersion Defined in ISO 12000 as liquids or semi-liquid materials, usually milky white, containing the polymeric material in a stable condition finely dispersed in a continuous liquid phase, normally water (aqueous dispersion).

Primer First coat applied to a substrate to (1) provide corrosion resistance, (2) improve adhesion to the substrate, and (3) provide specific performance (such as barrier coat for migrants from the substrate or to block gas permeation through plastics). A primer is of a much higher film weight than a size coat. A primer is a coating.

Resin Originally, natural macromolecular products derived from residues of trees, etc. Today, resins are mainly of synthetic origin, and those containing natural products, such as alkyds, need reaction with synthetic materials to improve their performance. Resins can be either thermoplastic or thermoset. All coatings contain resins.

Resin, thermoplastic These form film primarily through evaporation of solvent. Generally, they are higher molecular weights than thermoset resins.

Resin, thermoset (chemically cured) These form film through crosslinking reactions. Although defined as thermoset, application of heat is not necessary for reaction, with some systems curing at ambient. They are generally of low molecular weight and contain functional groups, such as hydroxyl, through which crosslinking occurs. A crosslinking agent is frequently present in the coating.

Size (size coat) Initially, size coats were used to seal porous substrates and provide adhesion for subsequent layers of coatings. Nowadays, size coats refer to low

film weight coatings applied over porous or non-porous substrates to improve the adhesion and mechanical and chemical resistance of subsequent coatings. A size coat is a coating.

Varnish External transparent (unpigmented) coating, used either for decoration or protection.

Vehicle Liquid component of a paint, that is, resins and solvents (i.e., a solution of the binder in the solvents).

I. HISTORY OF SURFACE COATINGS

The extent of the use of coatings, colorants, and paints can be used as an index of progress toward modern civilization. Primitive prehistoric paintings, such as those in the Altamira (Spain) and Lascaux (France) caves, were probably related to hopes of good hunting. The primitive painting of the grand bison on the walls of the Altamira cave in Spain is more than 15,000 years old. The painting of the Chinese horse on the walls of the cave at Lascaux, France, is less primitive than paintings at Altamira. The aborigines made pictures of game animals in their Obiri rock painting in Arnhem Land in northern Australia.

The first decorative coatings used on the walls of caves were based on colored mud. These water-dispersed coatings usually contained crushed berries, blood, eggs, plant sap, or milk. The Egyptians used distemper to decorate their walls in the period from 3000 to 1500 BC. The term *distemper* was coined by the British to describe waterborne coatings containing egg whites, glue, vegetable gums, or casein. The ancient Hebrews used milk or "curd paints" to decorate their homes.

The first paints were made by using binders (such as egg whites, milk curd, or pitch) to bind fillers and colorants (such as chalk, charcoal, ashes, earth colors, and vegetable dyes). The early Egyptians used pitch and balsam to caulk ships, and artists adopted these binders for paints. The first naturally based drying oil (unsaturated vegetable oil) was linseed oil obtained from flaxseeds.

The first colorants were white chalk, dark ashes, charred wood, and brighter hued earth colors. White lead [$2\text{PbCO}_3 \cdot \text{Pb}(\text{OH})_2$], the oldest synthetic pigment, was produced by Pliny, who placed lead sheets in vinegar in the presence of carbon dioxide more than 2500 years ago. This primitive pigment not only provided a white paint, but also accelerated the drying (curing) of vegetable-oil-based binders.

It has been reported (Mattiello, 1941) that there was a time when the surface coatings industry dealt with a relatively small number of materials and processes for making paints and varnishes. Indeed, it was claimed (Mattiello, 1941) that between 1736 and 1900, Watin's book on varnish formulations was reprinted 14 times with only minor modifications. This was claimed to be the in-

dustry standard (Mattiello, 1941). Compare that with today's situation. A book which lasted for 200 years to the turn of the century was followed by different books which have been superceded at decreasing time intervals. It is not the writing which is inadequate, but it is a true reflection of the increasing rate of change of the surface coating industry.

The early oleoresinous paints have been largely displaced by synthetic resins, either in solution, dispersion, or solvent free. A spirit (ethanol)-soluble phenolic resin (Bakelite) was introduced as a replacement for shellac at the beginning of the 20th century. Alkyd resin, essentially an oil-modified Glyptal resin, was introduced by R. H. Kienle in the 1920s. Vinyl acetate and vinyl chloride copolymers, aminoplasts, epoxy resins, and acrylic resins were introduced as paint resins during the 1930s–1950s. Paints based on aqueous polymeric emulsions and powder coatings were introduced in the 1950s. The last 10–20 years has seen a move toward more environmentally friendly coating systems, such as waterborne, high solids, radiation curable, and powder coatings.

II. NATURAL RESINS

Lacquers were introduced in the Chow dynasty (1027–256 BC) and were perfected in the Ming period (1368–1644 AD). The word *lacquer* is derived from *lac*, the name of a resin (shellac) secreted by an insect, *Coccus lacca*, that feeds on the lac tree. It has been estimated that 1,500,000 insects are required to produce sufficient lac for 0.5 kg of shellac (Mattiello, 1941). Shellac is the principal, non-vegetable, naturally occurring resin. The dried resin, called stick lac, is ground and washed to remove the red lac dye and is then melted and poured into molds. Shellac is a complex mixture of anhydrides, esters, and lactones of aliphatic and aromatic polyhydroxycarboxylic acids such as 9,10,16-trihydroxypalmitic acid. Because of its insolubility in varnish media, shellac has been used to cover knots in pine wood. Solutions of shellac are known as French polish for finishing wood. Shellac continues to be used in printing inks and as a spirit-soluble coating for candy, pills, and fruit.

However, shellac is no longer an important resin in the paint industry, due to cost and variability of supply.

Dammar resins are classified as recent fossil resins; that is, they are obtained from living plants and contain relatively large amounts of volatile solvents. These soft turpentine-soluble resins are obtained from trees in Borneo and Sumatra and are named after the export cities (i.e., Batavia, Singapore, or Penang). Dammar is used principally for its wax content.

Semi-fossil resins (i.e., those with relatively low solvent content) are obtained from dead trees. The principal East

Indies resins are batu and East India. These dark resins have been used in oleoresinous paints.

Melenket, loba, and Phillipine Manila resins are of recent origin, but pontianak is a semi-fossil resin and boea is a hard, oil-soluble fossil resin. Congo, a fossil resin from Africa, and kauri, a fossil resin from New Zealand, are also classified as copals.

Gum rosin (colophony) is obtained by tapping living pine trees or by the solvent extraction of pine tree stumps. The principal constituent of rosin is abietic acid, which can be isomerized by heat to produce levopimamic acid.

Gutta percha is a poly(trans-isoprene) obtained from the leaves of *Palaquium*, which is grown on plantations in Malaysia. Balata is a similar product obtained from the latex of *Minusops globosa*, which grows wild in Panama and northern South America. These resins have been used, to a limited extent, as paint resins but are no longer important in the coatings industry.

A naturally occurring fossil resin, amber, was used as an ingredient of varnishes as early as 250 BC. Actually, the name *varnish* is derived from *vernix*, the latin word for "amber." Amber is a fossil resin derived from an extinct variety of pine trees.

Accroides (Botany Bay gum), also called Accroides gum, is an exudate of the Australian grass tree (*Xanthorrhaea*). Elemi is a soft, yellow-brown recent resin obtained from *Canarium* and *Amyris* trees in the Phillipines. Sandarac gum (juniper gum) is a yellow recent resin obtained from *Calitris quadrivavis* trees in Morocco. Mastic gum is a turpentine-soluble exudate of the *Pistacia lentiscus* tree found in the Greek archipelago.

The first lacquers were produced from the exudate of the sumac tree (*Rhus vernicifera*).

III. OILS

A. Vegetable Oils

Vegetable oils are complex mixtures of triglycerides of fatty acids, ranging from about C₁₂ to C₂₀. Because climate, soil, and other natural factors affect the composition of oils, the same oil may vary in composition of its fatty acids from region to region and year to year. The most commonly used oils are C₁₈ fatty acids with different levels of unsaturation, typically none, one, two, or three double bonds. Apart from a few rare examples, these fatty acids possess unbranched chains and an even number of carbon atoms. When double bonds are present, they are usually in the *cis* form and are not conjugated.

Oils are classified into drying, non-drying, and semi-drying oils, depending upon the type and amount of unsaturation present in their fatty acid components. Examples of drying oils are linseed, tung (China wood), oticicia, and perilla. Examples of semi-drying oils are soya bean, tall

oil fatty acids, dehydrated castor, and safflower. Examples of non-drying oils are coconut, groundnut, olive, and hydrogenated castor.

Linseed oil has been used as a drying oil in coatings for more than eight centuries. Linseed oil, obtained from the seeds of flax (*Linum usitatissimum*) by expression or solvent extraction, consists of glycerides of linolenic, oleic, linoleic, and saturated fatty acids. This oil may be heated to 125°C with air (blown) to produce a drying oil that yields a harder film than unreacted linseed oil. "Boiled" linseed oil is oil that has been heated with traces of dryers.

Tung or China wood oil is obtained by roasting and pressing the seeds of *Aleurites cordata*. Tung oil contains the triple unsaturated eleostearic acid. The tung tree is indigenous to China and Japan, but has been grown in the Gulf Coast region of the United States.

Oiticica oil is obtained by pressing the seeds from the Brazilian oiticica tree (*Licania rigida*). Its principal constituents are glycerides of α -licanic acid (4-keto-9,11, 13-octadecatrienoic acid).

Dehydrated castor oil is obtained by heating oil from the castor bean (*Ricinus communis*) with sulfuric or phosphoric acid. During dehydration a double bond is formed.

Soybean oil is obtained by crushing the seeds of the legume *Glycine max*, *Soya hispida*, *S. japonica*, or *Phaseolus hispida*. This vegetable oil contains linoleic acid (51%), linolenic acid (9%), oleic acid (24%), and saturated oils (16%).

Tall oil fatty acids (TOFA) are derived from the stumps of pine trees. Scandinavian tall oil differs in composition to North American tall oil.

Coconut, palm, and groundnut oils are non-drying, primarily consisting of saturated fatty acids. They give good color retention to any system containing them.

B. Fish Oils

Menhaden oil is obtained from North Atlantic sardines (*Clupea menhaden*), and pilchard oil is obtained from European sardines (*C. pilchardus*). Oil from North American herring (*Clopea harengus*) is similar to sardine oil, and oil from Pacific herring (*C. caerulea*) resembles pilchard oil. About 70% of these fish oils are unsaturated oils. The use of fish oils has declined.

C. Air Drying of Oils

Drying oils such as linseed, soybean, and safflower oils have *cis*-methylene-interrupted unsaturation. Other oils, such as tung oil, contain conjugated double bonds.

The principal film-forming reaction of drying oils is oxidation, which includes isomerization, polymerization, and cleavage. These reactions are catalyzed by dryers

such as zirconium, cobalt, and manganese (organometallic salts).

It has been shown that oleic acid forms four different hydroperoxides at C-8, C-11, C-9, and C-10 and that unsaturated double bonds are present on C-9, C-10, and C-8. It is believed that these products undergo typical chain reaction polymerization via the standard steps of initiation, propagation, and termination.

Conjugated oils for use in oleoresinous paints have been produced by the dehydration of castor oil and the isomerization of linoleic and linolenic esters.

Linoleic and linolenic acids or esters can be converted to conjugated unsaturated acids or esters by heating with alkaline hydroxides or catalysts, which shift the methylene-interrupted polyunsaturation to conjugated unsaturations.

In addition to the dehydration of castor oil and isomerization of methylene-interrupted unsaturated oils such as linseed oil, conjugated succinyl adducts can be formed by the reaction of maleic anhydride with non-conjugated unsaturated oils. The succinyl adduct, which is attached at the 9 and 12 positions on the chain, undergoes rearrangement to produce a conjugated unsaturated acid. Additional rearrangement to a *trans-trans* form occurs at 200°C.

The reaction of 5–8% of maleic anhydride with soybean oil produces an unsaturated oil that is competitive with linseed oil and dehydrated castor oil.

D. Driers for Oils (and Alkyds)

Before 1840, white lead was used exclusively for the production of white oil paints. Initial attempts to replace the toxic lead pigment with zinc oxide were unsuccessful because the lead forms fatty acid salts, which act as dryers, whereas zinc salts are not primary dryers.

Oil-soluble salts of lead, manganese, and cobalt with fatty acids were introduced as dryers for oil-based paints in 1805. The most widely used acid for forming salts from dryers is now naphthenic acid, which is obtained in the refining of petroleum. Salts of tall oil from paper manufacture and salts of other metals such as vanadium, iron, zinc, calcium, and zirconium are also used. The effect of dryers is reinforced by the addition of 1,10-phenanthroline.

A combination of dryers is normally used. They can be classified into oxidation catalysts, polymerization catalysts, and auxiliary catalysts.

Oxidation catalysts, such as those based upon cobalt, manganese, cerium, vanadium, and iron(III) are primary dryers. They help the film absorb oxygen and participate in the formation and decomposition of peroxides. They greatly affect the surface hardness of the dried film.

Polymerization catalysts, based upon metals such as lead, zirconium, rare earths, and aluminium, assist drying to completion and are called through or secondary dryers.

Auxiliary catalysts based upon calcium, lithium, and potassium do not act alone as a dryer. They do, however, play a synergistic role with cobalt or zirconium. Calcium is the most widely used. Zinc is also in this category and is used to reduce the rate of surface drying, thereby permitting better oxygen absorption by the film as it dries.

IV. ESTER GUMS

Ester gum is produced by the esterification of rosin by glycerol. The ester obtained from pentaerythritol has a higher melting point than ester gum. Adducts of maleic anhydride and rosin are also used in varnishes and printing inks.

The most important maleic resin (ester gum) is obtained by the reaction of rosin with maleic anhydride followed by esterification with a polyol such as glycerol. Maleic acid resins are non-yellowing and can be used in oleoresinous varnishes or with alkyds.

V. OLEORESINOUS VARNISHES

Oleoresinous varnishes contain vegetable oils and are made by blending natural or synthetic resins with drying oils at high temperatures (500°F or more) and subsequently adding dryers and thinners. The original resins were fossil resins (copals) from various parts of the world. These resins were augmented by exudates of pine trees (balsams). Rosin is obtained when turpentine is distilled from the balsam.

Rapidly drying copolymers are produced by heating linseed oil with dicyclopentadiene at elevated temperatures. Syrenated oils are also obtained from linseed, soybean, dehydrated castor, and tung oils. There is some question as to whether copolymerization occurs in these reactions, but the products do give excellent films. Superior coatings are produced when the styrene or *p*-methylstyrene is grafted onto the unsaturated polymer chain.

VI. ALKYDS

Berzelius prepared glyceryl tartrate resins in 1847, and Smith prepared glyceryl phthalate resins (Glyptal) in 1901. Alkyds are a sub-class of polyesters, but due to their importance in coatings they are normally treated separately.

Patents were granted to Kienle for oil-modified polyesters in 1933. Kienle coined the euphonious name alkyd from the names of the reactants: alcohol and acid. Alkyds, the most widely used coating resins, are produced

by substituting some of the phthalic anhydride in the Glyptal formulation by an oil containing saturated or unsaturated fatty acids.

Oil-modified glycetyl phthalate resins are typical alkyd resins. The term *alkyd* also describes resins in which pentaerythritol, sorbitol, or mannitol has been substituted for glycerol. Likewise, while phthalic anhydride is normally used, isophthalic, terephthalic, succinic, adipic, azelaic, and sebacic acids as well as tetrachlorophthalic anhydride have been substituted for phthalic anhydride or phthalic acid in alkyd resins.

Alkyds are prepared by reacting oils with a polyol, such as glycerol, to form monoglycerides of the oil. These are then reacted with phthalic anhydride. Alternatively, fatty acids (derived from oils) can be reacted directly with polyols and phthalic anhydride.

Alkyds can be either air drying or non-air drying (normally baking), depending upon the type (conjugated or non-conjugated) and amount of carbon—carbon unsaturation present in the oils used. Different oils have different drying characteristics, and it is the combination of type of oil and oil length (medium to long) which determines whether an alkyd will air dry. Semi-drying oils can be used in air drying alkyd paints, but it depends upon the amount of oil present and the molecular weight of the alkyd.

Alkyds are classified as short-, medium-, and long-oil alkyds based on the relative amount of oil in the formulation. While definitions may vary slightly, short-oil alkyds typically contain <45% oil, medium-oil alkyds contain 45–55% oil, and long-oil alkyds contain >55% oil.

Non-drying alkyds can be used as a plasticizing resin in some coatings (e.g., nitrocellulose wood lacquers). In others, they need to be crosslinked to form a film (industrial stoving alkyd). Air drying alkyds have been used for many years in household gloss paints and have found extensive use.

A wide range of modified alkyds are available including styrenated, vinylated, and siliconized as well as urethane alkyds.

New generation alkyds, such as waterborne or high solids, will ensure the continued use of alkyd-based paints and coatings for many years to come.

VII. OIL-FREE POLYESTERS

Polyesters include alkyds. Normally, the differentiation used is alkyds and oil-free polyesters. The latter do not contain oils or fatty acids derived from oils. They consist of polyols and polyacids. Typical polyols include neopentyl glycol, trimethylol ethane (or propane), butane (or hexane) diol, and ethylene (or propylene) glycol. Typical polyacids include iso and terephthalic, adipic, and trimellitic anhydride. Phthalic anhydride finds limited usage in polyesters.

Polyester resins are used extensively in powder coatings and find use in certain types of industrial coatings. Siliconized polyesters are used for architectural claddings, coil applied.

Polyesters can contain unsaturated components such as maleic anhydride. These polyesters can be crosslinked by a free radical initiator. When dissolved in styrene or other unsaturated monomer, they find use in glass fiber laminates.

VIII. ACRYLIC RESINS AND STYRENE

Polyacrylic acid was synthesized in 1847, but this polymer and its esters were not investigated extensively until the early 1900s, when Otto Rohm wrote his Ph.D. dissertation on acrylic resins. Rohm and his associate, Otto Haas, produced acrylic esters commercially in Germany in 1927, and Haas brought this technology to the United States in the early 1930s.

Because the α -methyl group hinders segmental chain rotation, methacrylates are harder than their corresponding acrylate ester. Poly(methyl acrylate) has a T_g of -9°C compared to 105°C for poly(methyl methacrylate). The flexibility increases and T_g decreases as the size of the alkyl ester group increases. Poly(2-ethyl hexyl acrylate) has a T_g of -70°C , and poly(butyl methacrylate) has a T_g of 22°C .

Acrylic coatings are transparent and have good solvent, corrosion, and weather resistance.

Styrene is often copolymerized with acrylic monomers. This has technical benefits in some cases, as well as a cost reduction. Styrene has poor exterior durability characteristics, due to its aromatic ring. Polystyrenes are available, but find little usage as the sole coating resin. Styrene is nearly always copolymerized with acrylic monomers for use in coatings.

Acrylic monomers generally do not contain additional functionality. There are, however, a range of acrylate and methacrylate functional monomers available. This functionality includes hydroxyl, amino, amino-amide, acid, and epoxy. These monomers are incorporated into acrylic resins for thermoset applications or thermoplastic resins which may need the improved performance they bring, such as improved substrate adhesion. Thermoset acrylic coatings can be crosslinked by crosslinking agents, such as isocyanates, melamine, benzoguanamine, or urea formaldehyde resins. They may also be self-crosslinking as a result of functionality incorporated into the ester group. Acrylamide-containing resins are one example.

In 1956, General Motors Company replaced some of its cellulose nitrate automobile finishes with acrylic coatings. The labor-intensive polishing step used for both cellulose nitrate and acrylic finishes was eliminated by a “bake,

sand, bake" technique in which all sand scratches were filled in during the final bake. Thus, these coatings can replace cellulose nitrate and alkyd or melamine formaldehyde automobile finishes. Both solvent-based and water-borne acrylic coatings are available.

High-quality acrylic latices, typically made from methyl methacrylate and ethyl acrylate or related monomers, have found extensive use in water-based paints.

Acrylic resins find widespread usage in the coatings industry.

IX. PHENOLPLASTS

Phenol and substituted phenols can react with formaldehyde under basic or acidic conditions to form phenolic resins, sometimes called phenolplastics. Phenolic resins can be non-reactive (NOVOLACS) or reactive (RESOLES). Novolacs are the result of an excess of phenol under acidic conditions, while resoles are the result of an excess of formaldehyde under basic conditions. Novolac resins can participate in cross-linking reactions if a formaldehyde-generating material, such as hexamine, is added.

Phenolplastics are complex mixtures due to self-condensation, which occurs during their manufacture. Normally, the methylol groups are partially substituted by reaction with an alcohol, such as iso-butanol, to improve solubility, compatibility, and adjust reactivity.

The condensation products of phenol and formaldehyde developed by Leo Baekeland in 1909 were not suitable for oil-based coatings. However, condensates of formaldehyde with substituted phenols, such as *p*-octylphenol, are oil soluble and can be used with drying oils for making oleoresinous varnishes. Phenolic-modified oleoresinous spar varnishes have superior resistance properties, particularly water resistance.

Phenolic resins, per se, find limited usage in coatings, being primarily used for structural applications. Phenolic resins can be used as the sole binder in drum coatings, for example, but frequently other resins are present. Resoles are used to crosslink other resins, such as epoxy resins for can linings. Phenolic resins are used as binders for the sand in foundry moulds.

Substituted phenol-formaldehyde resins are also admixed with neoprene to produce industrial adhesives.

X. AMINOPLASTS

Urea, melamine, or benzoguanamine can react with formaldehyde to form amino resins sometimes called aminoplasts. Further reaction with monohydric alcohols produces etherified amino resins. They are complex mixtures and undergo some degree of self-condensation dur-

ing their manufacture. A major usage in coatings is as cross-linking agents.

Tollens condensed urea with formaldehyde in 1884. These resins are now used for paper coatings, wash-and-wear finishes on cotton textiles, and general-purpose coatings. The last-named are produced by the condensation of urea and formaldehyde in the presence of butanol. Butanol-etherified urea resins are used as modifiers for alkyd or cellulose nitrate resins in furniture coatings.

Melamine resins are produced by the condensation of melamine and larger proportions of formaldehyde than those used for urea resins. Like urea resins, they can be used as butylated resins. Also like the urea resins, they are used as crosslinking agents for alkyds, acrylics, epoxies, and other coating resins. Some melamine-based aminoplasts are water soluble, for example, hexamethoxyethyl melamine (HMMM).

Benzoguanamine formaldehyde resins are less functional and, hence, less reactive than their corresponding melamine formaldehyde resin.

XI. EPOXY RESINS

Epoxy resins, under the name of ethoxyline resins, were patented in Europe by P. Schlack in 1939 and introduced in the United States in the late 1940s. These resins, which are produced by the condensation of bisphenol A and epichlorohydrin, contain terminal epoxy groups and may contain many hydroxyl pendant groups, depending on molecular weight.

Epoxy resins are not used alone for coatings, normally being crosslinked.

Epoxy resins, based upon bisphenol A (or F) and epichlorohydrin, cured at room temperature by aliphatic polyfunctional amines and polyamides are used in heavy duty coatings for ships, oil rigs, and storage tanks, as well as water pipes. The epoxy resin-containing component of the paint is mixed with the polyamine-containing component prior to application. This is a two pack system. They also form the basis of the two pack adhesive (Araldite) available from most hardware stores.

Epoxy resins also react at elevated temperatures with aromatic amines, cyclic anhydrides, aminoplasts, and phenolplastics. In these applications, epoxy resin is present with the cross-linking agent in the coating and no mixing prior to application is required. This is a one pack coating. Epoxy resins have been used with acrylic coating resins, as powder coatings, and as high solids coatings. Epoxy resins are used as intermediates in UV coatings. Electrodeposition automobile primers are based upon epoxy resins crosslinked with polyisocyanate. Epoxy resins are used as additive resins in some coatings to improve properties such as adhesion and resistance.

Cycloaliphatic epoxy resins are used in cationic UV curable coatings and inks.

Epoxy resins reacted with fatty acids derived from oils produce epoxy esters, which can be air drying or non-air drying. The latter find usage in machinery coatings and are normally crosslinked with an aminoplast.

XII. POLYURETHANES

Urethane resins, introduced by Otto Bayer of I. G. Farbenindustrie in 1937, were produced in Germany during World War II. These polymers, which contain the characteristic NHCOO linkage, are produced by the condensation of a diisocyanate, such as tolylene diisocyanate, and a polyol, such as a polyether or polyester with two or more hydroxyl groups. Both aromatic and aliphatic isocyanates are used commercially, but great care must be exercised when handling them, particularly the more volatile isocyanates, due to their toxicity and potential for sensitization.

Polyurethanes can be produced *in situ* by reacting a diisocyanate and a polyether or polyester-containing hydroxyl group. These resins are thermoplastic and are used in inks for polyolefin films.

A major use of polyisocyanates is as a cross-linking agent for resins containing hydroxyl, or other, functionality. The resulting cross-linked network is a polyurethane. Because isocyanates are normally very reactive, they are sometimes used in two pack coatings.

One pack baked (stoving) coatings can be produced from blocked or capped isocyanates such as phenol adducts, which release diisocyanates when heated. These then react with any resins containing hydroxyl functionality.

One-component coatings based on the reaction products of drying oils and isocyanates are available. Monoglycerides are condensed with diisocyanates in a urethane-alkyd coating system. Moisture-cured polyurethane coatings are used for wood floor coatings.

Epoxy resins can be used as the polyol in cationic electrodeposited primers for automobiles. Polyurethane clear coats are used as an automobile topcoat.

XIII. VINYL ACETATE POLYMERS

Waterborne paints based on dispersions of casein were introduced commercially in the 1930s, but they lacked washing and scrub resistance. Hence, they were displaced by latices of polyvinyl acetate, stabilized by polyvinyl alcohol, and plasticized by dibutyl phthalate, which became available in the 1950s. More recent developments include

copolymers of vinyl acetate with ethylene and vinyl chloride to improve performance. Other soft comonomers, such as butyl acrylate, which act as internal plasticizers for the harder (higher glass-transition temperature T_g) homopolymer vinyl acetate, are also used.

As a general rule, vinyl acetate-based latices are restricted to internal usage, due to their poor exterior durability characteristics.

XIV. VINYL CHLORIDE AND VINYLIDENE CHLORIDE POLYMERS

A. Vinyl Chloride Containing Polymers

Polyvinyl chloride was synthesized by Baumann in the 1870s. However, because of its insolubility and intractability, it remained a laboratory curiosity until it was plasticized by Waldo Semon in the 1930s. In 1927 chemists at Carbon and Carbide, duPont, and I. G. Farbenindustrie produced copolymers of vinyl chloride and vinyl acetate that were more flexible and more readily soluble than polyvinyl chloride.

The most popular copolymer (Vinylite VYHH) contains 13% of the vinyl acetate comonomer and is soluble in aliphatic ketones and esters. A terpolymer (VYMCH) produced by the addition of maleic anhydride to the other monomers before polymerization has better adhesion than VYHH, and a partially hydrolyzed VYHH (VAGH) is more soluble than VYHH.

It is customary to add plasticizers and stabilizers to these coating resins. VAGH is compatible with alkyd, urea, and melamine resins. Copolymers with high vinyl acetate content (38SE, VYCC) are compatible with cellulose nitrate.

Polyvinyl chloride (PVC) suffers from dechlorination when heated, particularly in the presence of iron. Therefore, it is necessary to incorporate hydrochloric acid scavengers in PVC-containing coatings which will undergo thermal treatment, particularly if they are on metal. Examples of scavengers include epoxy resins and epoxidized oils.

B. Vinylidene Chloride-Containing Polymers

Polyvinylidene chloride was produced by Regnault in the 1830s, but this product was not commercialized until the 1930s. Chemists at Dow and Goodyear reduced the crystallinity of this polymer by copolymerizing the vinylidene chloride with acrylonitrile (Saran) or vinyl chloride (Pliovic). These copolymers, which are soluble in methyl ethyl ketone, are characterized by good resistance to gaseous diffusion and are used as barrier coatings. These copolymers are also used as latices for coating paper and textiles.

XV. CELLULOSE DERIVATIVES

Because of strong intermolecular hydrogen bonding, cellulose is insoluble in ordinary organic solvents. However, it is soluble in *N,N*-dimethylacetamide in the presence of lithium chloride, in dimethyl sulfoxide in the presence of formaldehyde, in Schweitzer's reagent (cuprammonium hydroxide), and in quaternary ammonium hydroxide. However, derivatives of cellulose are soluble in commonly used solvents.

A. Cellulose Nitrate—Nitrocellulose

Cellulose nitrate was synthesized in 1845 by Schonbein, who, because he believed it to be a nitro compound instead of an ester of nitric acid, mistakenly called it nitrocellulose. Solutions of cellulose nitrate (Pyroxylin) were patented by Wilson and Green in 1884.

Modern lacquers, which were introduced in 1925, are solutions of cellulose nitrate. These coatings dry by evaporation of the solvent. Pigmented nitrocellulose lacquers were used as finishes for automobiles in 1913.

Cellulose nitrate lacquers of higher solids content were hot-sprayed in the 1940s. In order to meet competition from alkyd-amino resin coatings, cellulose nitrate coatings were improved in the 1950s by the development of multicolor lacquer enamels and superlacquers based on cellulose nitrate-isocyanate prepolymers.

Wood lacquers were based upon nitrocellulose, but that usage is declining. Alkyd resins compatible with nitrocellulose have been used to plasticize nitrocellulose lacquers.

B. Cellulose Acetate

The first cellulose acetate was a triacetate produced by Schutzenberger in 1865. A more readily soluble cellulose diacetate was produced in 1903 by Miles, who reduced the acetyl content by partial saponification of the triacetate. The diacetate is soluble in acetone. More than 20 organic acid esters of cellulose are marketed in the United States, but only the nitrate and acetate are produced in large quantities. Cellulose acetate, plasticized by acetyl triethyl citrate, is used as a coating for paper and wire screening and as a cast film. Cellulose acetate butyrate (CAB) finds widespread usage as a modifying resin in industrial coatings.

C. Cellulose Ethers

Organosoluble ethylcellulose and ethylhydroxyethylcellulose are used in coatings and inks. These tough polymeric ethers have been approved by the U.S. Food and Drug Administration for use as food additives.

Methylcellulose, which is water soluble, was produced by Suida in 1905 and by Denham, Woodhouse, and Lilienfield in 1912. Water-soluble hydroxyethylcellulose was produced by Hubert in 1920, and carboxymethylcellulose was produced by Jansen in 1921. However, these polymeric ethers were not commercialized until the late 1930s, when they were used as thickeners in waterborne coatings systems.

XVI. RUBBER AND ITS DERIVATIVES

Natural rubber exists as a colloidal dispersion (latex) in certain shrubs and trees, the principal source being *Hevea brasiliensis*. In 1823, Charles Macintosh coated cloth with a solution of rubber in a hydrocarbon solvent to produce the rainwear that bears his name. Polychloroprene (Neoprene), produced commercially by duPont in 1932, was also used as a protective coating for cloth and metal.

Latices of natural rubber, styrene-butadiene rubber and acrylonitrile-butadiene rubber, have been used as coatings for textiles (e.g., rug backing), concrete, and steel. It is customary to add an aqueous dispersion of curing and stabilizing additives to the latex before application.

Chlorinated rubber containing ~67% chlorine and sold under the trade names Tornesit and Parlon has been used for several decades as a protective coating. Chlorinated rubber is insoluble in aliphatic hydrocarbons and alkanols, but is soluble in most other organic solvents. Chlorinated rubber coatings have good adhesion to Portland cement and metal surfaces and are resistant to many corrosive materials.

Cyclized rubber (Pliolite) is produced by the chlorostannic acid isomerization of natural rubber. Cyclized rubber is widely used in paper coatings. It is compatible with paraffin wax and many hydrocarbon resins. This resin is soluble in aliphatic and aromatic hydrocarbon solvents and is usually applied as a solution in toluene, xylene, or naphtha.

The trade name Pliolite S-5 is also used for a copolymer of styrene and butadiene. This resin is more economical and is often used in place of cyclized rubber.

XVII. SILICONES

Kipping synthesized polysiloxanes in the early 1900s, but failed to recognize their commercial potential. He believed these products to be ketones and, hence, called them silicones. These polymers, produced by the hydrolysis of dichlorodialkylsilanes, were commercialized in the 1930s by General Electric Co. and Dow Corning Co. Silicones have excellent resistance to water and elevated

temperatures. Aluminum powder-filled silicone coatings have been used at 250°C. Silicone-modified polyesters, alkyds, epoxy, and phenolic resins are also used as coatings. A modified silicone coating with a non-stick surface is used as a finish for other coating systems.

XVIII. POLYOLEFINS

Amorphous polyolefins, such as polyisobutylene, atactic polypropylene, and copolymers of ethylene and propylene, are soluble in aliphatic solvents, but crystalline polyolefin coatings must be applied as melts. Chlorosulfonated polyethylene produced by the reaction of polyethylene with chlorine and sulfur dioxide is soluble in aromatic solvents and is used as a protective coating. Because of the presence of the SO_2Cl pendant groups, this polymer can be cured by lead oxide in the presence of a trace of water or a diamine. The bonding of polyolefin coatings to metal substrates has been improved by the introduction of mercaptoester groups and by passing the polyolefin through a fluorine-containing gaseous atmosphere.

XIX. PIGMENTS AND COLORANTS

Pigments can be divided into organic and inorganic.

A. Inorganic Pigments

Basic lead carbonate [$2\text{PbCO}_3 \cdot \text{Pb}(\text{OH})_2$] was used by the early Greeks 2400 years ago, and other inorganic pigments such as zinc oxide (ZnO) and zinc sulfide (ZnS) were used in the late 1700s. A white pigment consisting of barium sulfate (BaSO_4) and zinc sulfate (ZnSO_4) (lithopone) was used widely in place of white lead. Ilmenite was used as a pigment by J. W. Ryland in 1865. Titanium dioxide, a white pigment, was extracted from this ore in the 1920s. It was synthesized in the 1950s and is now the most widely used white pigment. Lithopone and basic sulfate white lead pigments were introduced in the 1850s.

1. Titanium Dioxide

In 1908, Jebson and Farup produced white titanium dioxide (TiO_2) by the action of sulfuric acid on black ilmenite (FeTiO_3). In spite of its superior hiding power and because of its high price, this pigment was not used to any great extent until the 1930s.

In the past, the anatase form of the pigment was the most widely used, but now the more stable (exterior durability) rutile form predominates. The hiding power and tinting

strength of titanium dioxide is about twice that of other white pigments.

2. Iron Oxide

Earth pigments were used by the artists who painted on the walls of the Altamira caves in prehistoric times. These inexpensive pigments are still in use today.

a. Ochres. The coloring matter in ochres is exclusively iron oxide (Fe_2O_3). The pale golden French ochre contains from 20 to 50% iron oxide (Fe_2O_3). Yellow ochre, which is greenish yellow to brownish yellow, is hydrated iron oxide [$\text{FeO}(\text{OH})$]. The coloring material in red ochre is anhydrous ferric oxide. Burnt ochre is produced by the calcination of yellow ochre.

b. Sienna. This pigment was mined originally in Siena, Italy, but is now found throughout Italy and Sardinia. Raw sienna is yellow-brown and contains 50–70% iron oxide. Burnt sienna, which is reddish orange, is obtained by the calcination of raw sienna.

c. Umber. Raw umber, which is found in Cyprus, is greenish brown. It is contaminated with 12–20% manganese dioxide (MnO_2). Burnt umber, which is deep brown, is produced by the calcination of raw umber.

d. Other iron oxides. Haematite, which is found around the Persian Gulf, is α -ferric oxide, and this red pigment is called Persian red oxide. Limonite, a hydrated ferric oxide found in India and South Africa, is yellow. Magnetite (Fe_3O_4) is black. Many synthetic iron oxides are produced by the controlled calcination of copperas or green vitriol ($\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$).

3. Iron Cyanides

The generic name *iron blue* describes insoluble salts containing an iron hexacyanide anion. Prussian blue (ferrous ferricyanide) was discovered in Berlin in 1764. Prussian brown (ferric ferricyanide) is rapidly converted to Prussian blue by acidic surfaces such as clay.

4. Chromates

a. Chromic oxides. Guignet's green [$\text{CrO}(\text{OH})$] is obtained by heating an alkali metal bichromite with boric acid at 500°C. Chromium oxide green (Cr_2O_3) is obtained by the calcination of potassium dichromate in the presence of carbon.

b. Lead chromate. Chrome yellow is lead chromate (PbCrO_4). Lead oxide chromate ($\text{Pb}_2\text{O} \cdot \text{CrO}_4$) is orange-red. Actually, lead chromate may exist as a lemon yellow rhombic form, a reddish yellow monoclinic form, and a scarlet tetragonal form, but only the monoclinic form is stable at room temperature. Chrome orange and chrome red ($\text{PbCrO}_4 \cdot \text{PbO}$) are produced by heating lead chromate with alkali. The difference in the two colors is related to particle size. Chrome green is a mixture of chrome yellow and Prussian blue. Zinc yellow is zinc chromate loosely combined with zinc hydroxide.

5. Lead Pigments

Basic sulfate white lead ($\text{PbSO}_4 \cdot \text{PbO}$) can be produced by vaporizing galena (lead sulfide, PbS) and burning the vapor with a controlled amount of air or by spraying molten lead into a furnace with sulfur dioxide and air. Leaded zinc oxide is a mixture of white lead and zinc oxide. Lead tungstate (PbWO_4) and lead vanadate [$\text{Pb}(\text{VO}_3)_2$] are yellow pigments. Because all lead compounds are extremely poisonous, their use is being discontinued.

6. Zinc Pigments

a. Zinc oxide. Zinc oxide (ZnO) has been used as a white pigment since 1835. Zinc oxide occurs in the United States as franklinite, willemite, and zincite ores, which are roasted to produce zinc oxide. The oxide is reduced to zinc by heating with carbon. In both the synthetic and mineral processes, the zinc metal is volatilized and oxidized by air to zinc oxide. On a weight basis, the hiding power of zinc oxide is superior to that of white lead. It acts as a moldicide and an ultraviolet absorber.

b. Lithopone and zinc sulfide. Zinc sulfide (ZnS) is used directly as a white pigment and also with barium sulfate (BaSO_4) as a component of lithopone. Lithopone, which is used as a white pigment in waterborne paints, has better hiding power than zinc oxide.

7. Antimony Pigments

The white pigment antimony oxide (Sb_2O_3), sold under the trade name Timonox, has been used to reduce chalking of anatase-pigmented paints. However, it is not needed when the rutile form of titanium dioxide is used. The principal use of antimony oxide is as a flame retardant.

8. Cadmium Sulfide

Cadmium sulfide (CdS), called cadmium yellow, lacks acid resistance but is acceptable for indoor use in neutral or alkaline environments. This pigment tends to react with

carbon dioxide and form less yellow cadmium carbonate. Cadmium sulfide can be admixed with barium sulfate to produce cadmopone. Because cadmium compounds are extremely toxic, their use is being discontinued.

9. Cobalt Blue

Because cobalt blue ($\text{CoO} \cdot \text{Al}_2\text{O}_3$) was introduced as a pigment by Thenard and Proust in the late 1800s, it is often called Thenard's blue. Because of its low refractive index, it is transparent in oily media. It is used as a blue pigment in ceramics, but is not competitive with phthalocyanine blue in coatings.

10. Ultramarine

The ancient Egyptians obtained natural ultramarine by pulverizing the precious stone lapis lazuli. The name, which means "beyond the sea," refers to the asiatic origin of this semiprecious stone. In 1828, Guimet and Gmelin independently synthesized this pigment in their attempts to win a 6000 franc prize from the Société d'Encouragement in Paris.

In the 1870s, more than 30 plants were producing synthetic ultramarine by heating clay, sulfur, and sodium sulfate with charcoal. Through further treatment, one obtains ultramarine in shades ranging from green to red. Ultramarine is used as a pigment for paper, textiles, plastics, and rubber.

11. Metallic Pigments

Finely divided metals and alloys are used as pigments in paints. The principal pigments are obtained from aluminum, zinc, and bronze. Aluminum powder, in both leafing and non-leaving forms, is produced by atomization of molten aluminum. The atomized powder is converted to lamellar particles by ball milling with stearic acid.

Metallic powders and lamellar pigments are produced from bronze by techniques similar to those employed for aluminum. These pigments are used for decorative effects. Zinc dust is used in the manufacture of anticorrosive primers. Metallic flakes are used for the production of antistatic coatings.

B. Organic Pigments

Natural organic colorants, such as carmine, logwood, indigo, and madder, which were used by ancient people, have been replaced by synthetic pigments and dyes. The first synthetic dyestuff was prepared in 1856 by Perkin, who oxidized aniline.

The search for these synthetic colorants was catalyzed in 1876 by Witt, who identified azo ($-\text{N}=\text{N}-$) and other

common chromophores. Organic pigment dyestuffs are insoluble, while soluble toners are precipitated by the addition of metallic salts.

1. Carbon Black

The first carbon black, called lamp black, was produced 5000 years ago by allowing a flame to impinge on the surface of an inverted bowl. More recently, bone black has been prepared by the carbonization of degreased bones. Carbon black is now produced in large quantities by the incomplete combustion of hydrocarbons. The tinting strength of carbon black pigment is inversely proportional to its particle size.

Channel black has a "blackness or jetness" that is superior to that of furnace blacks. The latter, whose production is more acceptable to environmentalists, is produced by the incomplete combustion of natural gas or oils in a special furnace. Channel black is produced by an impingement process similar to that used to produce lamp black. Furnace blacks are used as reinforcing agents in rubber, but as pigments they are inferior to other blacks.

2. Aniline Black

Aniline black (nigrosine), which is obtained by the oxidation of aniline in the presence of Cu(II) ions, is the most widely used organic pigment. Unlike carbon black, the hiding power of aniline black is due to absorption rather than the scattering of light.

3. Phthalocyanines

Phthalocyanines, which were synthesized in 1927 by de Diesbach and van der Weid, have to a large extent replaced alkali-sensitive iron blues and acid-sensitive ultramarine. Gaertner has called the discovery of phthalocyanines "the turning point in pigment chemistry."

This pigment is produced by dissolving the heterocyclic phthalocyanine in concentrated sulfuric acid and precipitating it by the addition of water. Copper phthalocyanines are stable to heat and most corrosives. Green phthalocyanine is obtained by the chlorination of blue phthalocyanine. Bromination produces a yellow product.

4. Azo Colors

Insoluble azo pigments are produced by coupling diazonium compounds with a phenolic compound such as 2-naphthol, acetoacetanilide, or pyrazolone derivatives. Soluble azo pigments, which are characterized by the presence of acidic groups, are precipitated from aqueous solution by barium, calcium, iron, copper, or manganese

cations. Typical azo pigments are toluidine red, Hansa yellow, benzidine yellow, pyrazolone red, and the barium salts of lithol red and pigment scarlet.

5. Vat Pigments

Vat pigments are an important class of textile colorants. Naturally occurring indigo and madder were known to the earliest civilizations but are made synthetically today. The name *vat dye* is derived from the vat in which the reduction of color and the immersion of cloth occurs. Reoxidation takes place in the atmosphere after removal from the vat.

6. Quinacridones

Quinacridone and its derivatives are stable solvent-resistant pigments. The color of these pigments ranges from red-yellow to violet, scarlet, and maroon.

XX. SOLVENTS

Ethanol, which was used for dissolving shellac, was produced a few thousand years ago by the fermentation of starch from grains. Ethyl ether, which was produced in the 16th century by the reaction of sulfuric acid and ethanol, was called sulfuric ether. It was perhaps fortuitous that a mixture of the few available solvents, such as ethyl ether and ethanol, was used to produce a solution of cellulose nitrate (collodion) in the middle of the 19th century.

Turpentine was produced by the distillation of gum from pine trees in the 18th century and became the solvent of choice for the paint industry. The general use of ethanol as a solvent was delayed until 1906, when tax-free denatured alcohol became available.

Weizman produced acetone and butanol by the bacillus fermentation of corn during World War I. Methanol, produced by the destructive distillation of wood, has been available for more than a century, but this process has been replaced by the conversion of synthesis gas (CO and H₂) to methanol. Isopropyl alcohol, produced by the oxidation of propylene, has been available commercially since 1917. Aromatic hydrocarbon solvents, toluene and naphtha, were introduced in 1925. There is a trend away from some aromatic solvents toward aliphatic ones for toxicity reasons.

Oxygenated solvents became available about 60 years ago to meet the solvency requirements of emerging synthetic resins and some industrial coating application processes. Some of the early oxygenated solvents were shown to possess adverse health effects, some causing abnormalities in the unborn child. A modern generation of

oxygenated solvents have been developed to minimize health risks.

In spite of the wide variety of solvents now available, the use of solvents in coatings and paints is decreasing because of restrictions on their use, as recommended by regulatory agencies such as the U.S. Environmental Protection Agency and European Union Directives. Hence, the trend is toward the use of coatings with a higher solids content (less solvent), waterborne coatings, and radiation curable and powder coatings.

XXI. RADIATION CURABLE COATINGS

The use of electromagnetic radiation to crosslink or cure polymeric coatings has been called radiation curing. The applied radiation may be ionizing radiation (i.e., α , β , or γ rays) or high-energy electrons and non-ionizing radiation, such as UV, visible, IR, microwave, and radiofrequency wavelengths of energy.

Most workers refer to radiation curable coatings as those which cure by exposure to UV radiation or an electron beam. There are two distinct types of radiation curable coatings. One is a free radical mechanism utilizing acrylic unsaturation. The other is cationic. Both are characterized by being solvent free, so they meet solvent emissions and any potential carbon dioxide taxes.

Relatively low molecular weight oligomers normally containing acrylic unsaturation are normally dissolved in "monomers" containing acrylic unsaturation. The oligomers are often reaction products of acrylic acid with epoxy resins, hydroxy acrylates with polyisocyanates (or a low molecular weight polyester urethane containing residual isocyanate functionality), or a low molecular weight polyester reacted with an acrylic monomer, such as acrylic acid. Methacrylate unsaturation is not normally used, because it is too unreactive at the speeds at which modern equipment runs. The monomers are either multifunctional or monofunctional. Most are multifunctional and are obtained by reacting a polyol with acrylic acid, for example, TPGDA—tripropylene glycol diacrylate.

The "wet" coating contains a photoinitiator package. This may be one or more materials which decompose on exposure to UV radiation of certain wavelengths generating free radicals. Synergists may also be present to improve free radical generation of the photoinitiator, and some synergists also reduce oxygen inhibition of curing. Different photoinitiators are essential, partly to overcome pigments absorbing the UV radiation at their absorption spectrum. A mixture of photoinitiators with different absorption spectra can be used. The free radicals, once formed, initiate polymerization of the coating film through the unsaturation in the monomers and oligomers.

The light source for UV curable coatings is a medium-pressure mercury arc lamp enclosed in quartz. In the last 30 years or so, much progress has been made with the technology and chemistry. Radiation curable coatings find widespread usage for paper and board, plastics, optical fibers, digital optical recording and laser vision discs, adhesives, dental fillings, circuit board coatings, and wood. There are limitations for metal at present. UV curable inks are an important area for this technology.

Electron beam cured coatings are similar to the UV ones mentioned above, but they do not contain a photoinitiator. The electron beam has sufficiently high energy to fragment the molecules in the coating and initiate polymerization. There are only a few installations. The capital cost of the equipment is high, and, due to oxygen inhibition, an inert gas blanket is required.

Cationic UV curable coatings, which have been commercial since about 1980, contain monomers, oligomers, and photoinitiators. Typically, the monomers are vinyl or propenyl ethers and the oligomers are aliphatic epoxies. The photoinitiators, often aryl sulphide or phosphonium halogen salts, degrade on exposure to UV radiation. Cationic coatings do not suffer from oxygen inhibition of cure, but moisture in the atmosphere is a major problem. Cationic coatings have good performance on metals. The penetration of UV cationic coatings is less than some imagined when this technology was originally commercialized.

XXII. HIGH SOLIDS COATINGS

The solvent content of many classical coatings has been reduced in order to comply with environmental protection regulations. The first solvent emission law (Rule 66) was enacted in California in 1966. Subsequent regulations on state and national levels have resulted in a reduction of solvent contents. Europe also has controls on solvent emissions. High solids coatings refer to those coatings which contain solvent, or volatile materials, but at a much reduced level than their equivalents a few years ago. There are different amounts of solvents in different types of coatings which are all classified as high solids. A particular industry and geographic location may have different definitions to the same industry elsewhere or to other industries in the same location. As a general rule, high solids is not used by industrialists to refer to solvent-free coatings.

Application viscosity is a limiting parameter for coatings. As a general rule, the higher the molecular weight of a resin, the greater the solution viscosity. Therefore, lower molecular weight resins and reactive diluents have been developed. Too low a molecular weight of the resin will result in poor film performance, so a compromise is

required. There have been many developments directed toward retaining high molecular weights, while keeping the viscosity relatively low. The other development has been toward reactive diluents with improved performance characteristics.

Thermoset coatings containing relatively low molecular weight resins are ideal candidates for high solids, particularly if a high solids aminoplast is used as the cross-linking agent. It is difficult to obtain high solid solutions of high molecular weight thermoplastic resins, with dispersion being one of the few viable routes.

High solids alkyds, polyesters, acrylics, aminoplasts, and other resins are commercially available. If solvent content cannot be reduced, then alternatives, such as waterborne, are often used.

XXIII. POWDER COATINGS

The first powder coatings were based on low-density polyethylene, which is too insoluble to be used as a solution coating. The first powder-coating process was based on flame spraying, and this was replaced by a process in which the article to be coated was heated and dipped in a fluidized bed of powdered polymer. The loosely coated article was then placed in an oven in order to sinter the powder and obtain a continuous film.

Thinner coatings of thermosetting resins, such as epoxy resins, are now applied by an electrostatic spray gun and then cured by heat. Epoxy, acrylic, polyester, and polyurethane chemistries, as well as hybrids of some, all find use in powder coatings. The majority of powder coatings are thermoset systems, thus the cross-linking agents must also be a solid at room temperatures. A sharp melting point is a prerequisite for a good flowing powder coating to give good gloss and appearance. The melting point has to be well above ambient to ensure that the powder particles do not sinter during storage before application.

The use of electrostatic spraying and the need to store above 100°C (often 150°C+) make powder coatings ideally suited to coating metal articles. Therefore, it is not surprising that this technology has replaced traditional wet coating chemistries in many industrial applications, such as “white goods.”

Substrates must be able to withstand the temperatures necessary for flow and cure, although newer developments are aimed at overcoming these deficiencies. An example would be the emerging generation of powder coatings for wood.

XXIV. WATERBORNE COATINGS

Waterborne coatings were used by the Egyptians 5000 years ago, and “distemper,” which refers to waterborne

coatings containing casein, glues, and other components, continues to be used as artists’ colors. The white wash used by Tom Sawyer in Mark Twain’s book was water-slaked lime. The adhesion and durability of this dispersion of lime were improved by the addition of casein.

Pigmented casein paints were introduced in the 1930s. An improved film deposit was obtained by the addition of a drying oil, such as linseed oil, and brightly colored casein paints were widely used at the Century of Progress Exposition in Chicago in 1933.

The use of casein-dispersed waterborne coatings in the 1940s was catalyzed by the introduction of roller coating techniques in “do it yourself” applications. Polyvinyl acetate, styrene-butadiene copolymer, and acrylic latices were introduced after World War II.

Waterborne coatings cover many categories of resin. There are only a few truly water-soluble resins used in coatings. Some cross-linkers, such as HMMM, are truly water soluble. One of the drawbacks of being water soluble is a sensitivity to moisture and poor water resistance properties of the dried paint film. Most waterborne coatings are derived from water-insoluble resins, by either polymerizing their monomers in water (polyvinylacetate, styrene-butadiene copolymers, and acrylics latices) or preparing a resin with either acid or basic functionality, neutralizing this functionality and dispersing in water. This process is sometimes called inverting a resin (from solvent based to water dispersed). These coatings may be (1) anionic (i.e., have a neutralizable functional carboxyl group), (2) cationic (i.e., have a neutralizable functional amine group), or (3) non-ionic (i.e., have hydroxyl solubilizing groups). During curing or the drying process, the neutralizing agent leaves the film, thereby restoring the water insolubility of the original resin. Most of the dispersed waterborne resin systems contain water-soluble solvents (butyl glycol, for example), sometimes referred to as cosolvents. Dispersed pigments and additives, such as defoamers, flow aids, emulsifiers, and coalescing agents (for emulsions), are usually present in waterborne formulations.

It is possible to emulsify other resins by dispersing in water using emulsifiers. Emulsions of epoxy resin and curing agents are used for waterborne concrete coatings and some heavy duty applications, where solvent fumes would be hazardous.

The electrodeposition process, in which charged paint particles are electrically plated out of water suspensions onto a metal article, was developed by Crossee and Blackweld in England in the 1930s and used by Brewer to coat Ford automobile frames in 1963. Electrodeposited coatings are theoretically of uniform thickness, regardless of the shape and complexity of the substrate. They can penetrate areas which would be very difficult, if not impossible, to coat by any other means, such as automobile chassis.

Originally, most water-soluble coatings used in electrodeposition were anionic, having carboxyl groups in the polymer chain, which were neutralized by the addition of sodium hydroxide or an organic amine. Some of the resins used were maleinized and styrenated drying oils, alkyds, polyesters, and acrylics.

Cationic electrodeposition has largely superceded anionic electrodeposition, mainly because better corrosion resistance is obtained. Automobile primers are based upon epoxy resin reacted with amines and neutralized by acid. Frequently, these coatings are crosslinked by a polyisocyanate.

Other applications for electrodeposited paints include central heating radiators and industrial metallic parts such as shelving.

There is a continuing growth in waterborne coatings as many industries are moving away from solvent-borne coatings.

XXV. TOXICITY

Some solvents, pigments, and binder components used in coatings are injurious to human organisms, and, hence, appropriate precautions must be taken to prevent accidental contact with these toxic ingredients. Extreme care must be used in the application of coatings containing lead, cadmium, or chrome(VI) pigments. The use of these pigments has substantially decreased. Inhalation of organic solvents, such as toluene and chlorinated solvents, must be avoided during the application of coatings. Extraction or well-ventilated areas must be used during the application of a coating containing solvent. Likewise, care must be exercised when polyurethane containing volatile isocyanates and epoxy resins containing free amines are applied. Sensitization by epoxies and isocyanates is another problem, and skin contact should be avoided with any coating. Solvent-containing paints must be applied in a well-ventilated, flame-free area.

XXVI. FUTURE TRENDS

The primary functions of paint are decoration and protection. Neither of these qualities is required for pigmented plastics and inherent decorated surfaces such as polyvinyl chloride siding and composite wallboard.

The high capital cost of cars, equipment, ships, and storage vessels necessitates their adequate protection. This can only be achieved by the application of a coating. In some other cases, such as food cans, the coating is not only protective, but is essential for the manufacture and filling of the can.

The use of traditional paints and coatings containing high levels of noxious solvents will continue to reduce.

Replacement paints and coatings will be more environmentally and user friendly. Aliphatic solvents will replace aromatic ones wherever possible. The use of toxic pigments will continue to decrease. People applying coatings and the end consumers will become more aware of any environmental or health problems associated with those coatings and changes will be enforced to overcome them.

The penetration of waterborne, powder, and high solids coatings will continue. It is probable that the move toward radiation curable inks and coatings will slow down, due to the fact that most applications suited to this technology have converted to it, with further technological developments being required before further inroads are made.

The proposals for carbon dioxide taxes in some European countries will drive the market. The increasing costs of energy will drive thermally cured coatings to lower cure schedules. Economics will also play a major role, and coatings and coated articles must remain competitive with other forms of materials, for example, laminated metal.

Perhaps the biggest change will be in those countries (many being considered as Third World) which have largely not moved from traditional coatings. They will eventually move to safer and more environmentally coatings. The only question is when?

SEE ALSO THE FOLLOWING ARTICLES

BIOPOLYMERS • COLOR SCIENCE • KINETICS (CHEMISTRY) • POLYMERS, SYNTHESIS • RUBBER, NATURAL • RUBBER, SYNTHETIC • SURFACE CHEMISTRY

BIBLIOGRAPHY

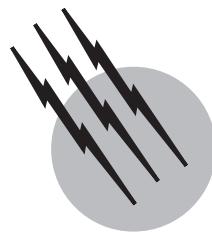
History of surface coatings

- Mattiello, J. J. (1941). "Protective and Decorative Coatings," Vol. I, Wiley, New York.
- Mattiello, J. J. (1941). "Protective and Decorative Coatings," Vol. II, Wiley, New York.
- Mattiello, J. J. (1943). "Protective and Decorative Coatings," Vol. III, Wiley, New York.
- Mattiello, J. J. (1941). "Protective and Decorative Coatings," Vol. IV, Wiley, New York.
- Mattiello, J. J. (1946). "Protective and Decorative Coatings," Vol. V, Wiley, New York.

More detailed and recent treatises

- Bates, D. A. (1990). "Powder Coatings," SITA Technology, ISBN 0947798005.
- Holman, R., and Oldring, P. K. T. (1988). "UV & EB Curing Formulation for Printing Inks, Coatings & Paints," SITA Technology, ISBN 0947798021.
- Landbourne, R., ed. (1987). "Paint and Surface Coatings: Theory and Practice," Ellis Horwood, Chichester, UK.

- Lee, L. H., ed. (1988). "Adhesives, Sealants, and Coatings," Plenum, New York.
- Lowe, C., and Oldring, P. K. T. (1994). "Test Method for UV & EB Curable Systems," SITA Technology, ISBN 0947798072.
- Martens, C. R. (1981). "Water Borne Coatings," Van Nostrand-Reinhold, New York.
- Morgans, W. M. (1984). "Outlines of Paint Technology," Vols. I and 2, Charles Griffin and Co., High Wycombe, UK.
- Oldring, P. K. T., ed. (1986). "Resins for Surface Coatings," Vol. I, SITA Technology, ISBN 0947798048.
- Oldring, P. K. T., ed. (1987). "Resins for Surface Coatings," Vol. II, SITA Technology, ISBN 0947798056.
- Oldring, P. K. T., ed. (1987). "Resins for Surface Coatings," Vol. III, SITA Technology, ISBN 0947798064.
- Oldring, P. K. T., ed. (1991). "Chemistry & Technology of UV & EB Formulation for Coatings, Inks & Paints," Vol. I, SITA Technology, ISBN 0947798110.
- Oldring, P. K. T., ed. (1991). "Chemistry & Technology of UV & EB Formulation for Coatings, Inks & Paints," Vol. II, SITA Technology, ISBN 0947798102.
- Oldring, P. K. T., ed. (1991). "Chemistry & Technology of UV & EB Formulation for Coatings, Inks & Paints," Vol. III, SITA Technology, ISBN 0947798161.
- Oldring, P. K. T., ed. (1991). "Chemistry & Technology of UV & EB Formulation for Coatings, Inks & Paints," Vol. IV, SITA Technology, ISBN 0947798218.
- Oldring, P. K. T., ed. (1994). "Chemistry & Technology of UV & EB Formulation for Coatings, Inks & Paints," Vol. V, SITA Technology, ISBN 0947798374.
- Oldring, P. K. T., and Lam, P., eds. (1996). "Waterborne & Solvent Based Acrylics and Their End User Applications," SITA Technology, ISBN 0947798447.
- Oldring, P. K. T., ed. (1997). "Waterborne & Solvent Based Epoxies and Their End User Applications," Wiley, New York.
- Oldring, P. K. T., ed. (1999). "The Chemistry and Application of Amino Crosslinking Agents or Aminoplasts," Wiley, New York.
- Oldring, P. K. T., ed. (1999). "The Chemistry and Application of Phenolic Crosslinking Agents or Phenoplasts," Wiley, New York.
- Sanders, D., ed. (1999). "Waterborne & Solvent Saturated Polyesters and Their end User Applications," Wiley, New York.
- Seymour, R. B. (1982). "Plastics vs Corrosives," Wiley, New York.
- Seymour, R. B., and Carraher, C. E. (1988). "Polymer Chemistry: An Introduction," 2nd ed., Dekker, New York.
- Seymour, R. B. (1988, 1989, 1990). "Advances in coatings science technology," *J. Coat Technol.* **60** (7S9), 57; **61** (776), 73; **62** (781), 62, 63.
- Seymour, R. B., and Mark, H. F., eds. (1990). "Organic Coatings: Their Origin and Development," Elsevier, Amsterdam, The Netherlands.
- Seymour, R. B., and Mark, H. F., eds. (1990). "Handbook of Organic Coatings," Elsevier, Amsterdam, The Netherlands.
- Solomon, D. H., and Hawthorne, D. G. (1983). "Chemistry of Pigments and Fillers," Wiley, New York.
- Tess, R. W., and Poehlein, C. W., eds. (1985). "Applied Polymer Science," 2nd ed., Am. Chem. Soc., Washington, DC.
- Thomas, P., ed. (1998). "Waterborne & Solvent Based Polyurethanes and Their end User Applications," Wiley, New York.



Composite Materials

M. Knight

D. Curliss

Air Force Research Laboratory

- I. Characteristics
- II. Constituent Materials
- III. Properties of Composites
- IV. Analysis of Composites
- V. Fabrication of Composites
- VI. Uses of Composites

GLOSSARY

Advanced composites Composite materials applicable to aerospace construction and consisting of a high-strength, high-modulus fiber system embedded in an essentially homogeneous matrix.

Anisotropic Not isotropic; having mechanical and/or physical properties that vary with direction relative to a natural reference axis inherent in the materials.

Balanced laminate Composite laminate in which all laminae at angles other than 0° and 90° occur only in ±pairs.

Constituent In general, an element of a larger grouping. In advanced composites, the principal constituents are the fibers and the matrix.

Cure To change the properties of a thermosetting resin irreversibly by chemical reaction.

Fiber Single homogeneous strand of material, essentially one-dimensional in the macrobehavior sense.

Interface Boundary between the individual, physically distinguishable constituents of a composite.

Isotropic Having uniform properties in all directions. The measured properties are independent of the axis of testing.

Lamina Single ply or layer in a laminate made of a series of layers.

Laminate Unit made by bonding together two or more layers or laminae of materials.

Matrix Essentially homogeneous material in which the reinforcement system of a composite is embedded.

Orthotropic Having three mutually perpendicular planes of elastic symmetry.

Transversely isotropic Material having identical properties along any direction in a transverse plane.

Woven fabric composite Form of composite in which the reinforcement consists of woven fabric.

1, or x, axis Axis in the plane of the laminate that is used as the 0° reference for designating the angle of a lamina.

2, or y, axis Axis in the plane of the laminate that is perpendicular to the x axis.

3, or z, axis Reference axis normal to the plane of the laminate x, y axes.

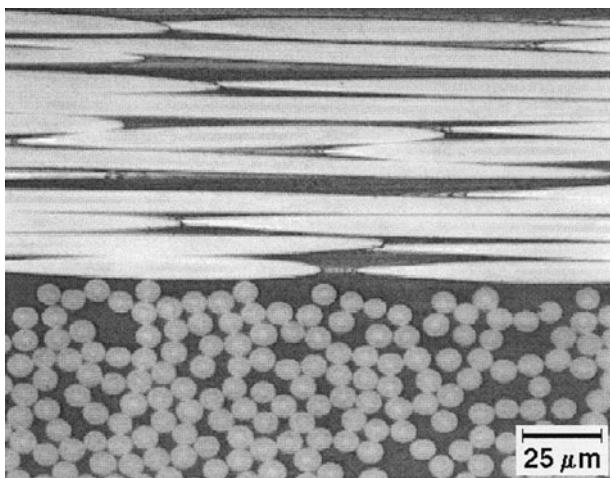


FIGURE 1 Cross section of a graphite fiber-reinforced epoxy polymer.

A COMPOSITE MATERIAL is described in this chapter as a material composed of two or more distinct phases and the interfaces between them. At a macroscopic scale, the phases are indistinguishable, but at some microscopic scales, the phases are clearly separate, and each phase exhibits the characteristics of the pure material. In this chapter, we are only describing the characteristics, analysis, and processing of high-performance structural composite materials. This special class of composites always consists of a reinforcing phase and a matrix phase. The reinforcing phase is typically a graphite, glass, ceramic, or polymer fiber, and the matrix is typically a polymer, but may also be ceramic or metal. The fibers provide strength and stiffness to the composite component, while the matrix serves to bind the reinforcements together, distribute mechanical loads through the part, provide a means to process the material into a net shape part, and provide the primary environmental resistance of the composite component. In Fig. 1, we can see the distinct cross section of graphite fibers in an epoxy matrix.

I. CHARACTERISTICS

Many materials can be classified as composites. They are composed of several distinctly different and microscopically identifiable substances. Composites are widely used in many industries and applications today, driven by the need for strong, lightweight materials. The composites reduce weight and allow for designs that tailor the mechanical properties of the material to meet the loading requirements of the structure. In addition, composites are replacing traditional engineering materials in many industrial, recreational, architectural, transportation, and infrastructure applications.

Composites occur very commonly in nature. Some of the best examples are wood, bone, various minerals, mollusk shells, and insect exoskeletons. In wood, the cellulose fibers of the cell wall are “glued” together by the lignin matrix. Bone is composed of calcium hydroxyapatite crystals in a protein matrix. Mollusk shells are composites of calcium carbonate layers in various geometries bound together by a multilayer matrix. Insect exoskeletons bear a striking resemblance to man-made fiber-reinforced composites. Some insects even exhibit apparent “layers” of fibrous chitin embedded in a protein matrix, where the orientation of the fibers varies from layer to layer, much as we might design a man-made fiber-reinforced composite. This example of a natural composite can be clearly seen in Fig. 2. Modern materials engineers have used the composite concept—reinforcement in a matrix—to create a class of modern materials that offers opportunities significantly greater than those of more common engineering materials.

Composites can be made of a such a wide variety of materials that it is impractical to discuss each one individually. One of the principal characteristics of all composites is that they have a reinforcement phase distinct from the matrix phase. The individual characteristics of the two phases combine to give the composite its unique properties.

Classes of materials commonly used for reinforcements are glasses, metals, polymers, ceramics, and graphite. The reinforcement can be in many forms, such as continuous fibers or filaments, chopped fibers, woven fibers or yarns, particles, or ribbons. The criteria for selecting the type and form of reinforcement vary in accordance with the design requirement for the composite. However, certain general qualities are desirable, including high strength, high modulus, light weight, environmental resistance, good elongation, low cost, good handleability, and ease of manufacture. By far, the most widely used reinforcement is E-glass.

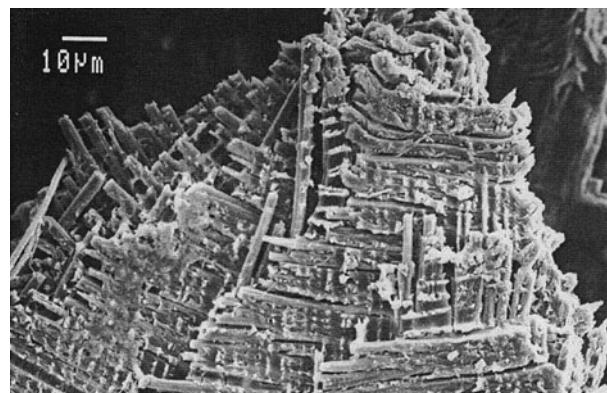


FIGURE 2 Scanning electron microscope (SEM) image of a bessbeetle (*Odontotaenius disjunctus*) elytra fracture surface.

E-glass offers excellent strength, compatibility with common matrix polymers, and is very low in cost. Various types of graphite fibers are commonly used in aerospace and the recreational products industry, where light weight and maximum material performance are very important to the designer.

The matrix binds the reinforcement together and enhances the distribution of the applied load within the composite. Polymeric materials are widely used as matrix materials. Two general classes of polymers are used: thermosets and thermoplastics. Thermosets are initially low molecular weight molecules that are often viscous liquids at room temperature—what we commonly think of as “resins.” Their low viscosity and fluid behavior make them very suitable to low-cost processing. The thermoset resins undergo chemical reactions when heated (or initiated by some other energy source such as UV light, electron beam, or microwave) and form a high molecular weight cross-linked polymer. In contrast, thermoplastics are high molecular weight linear polymers that are fully formed prior to processing as a composite matrix. When heated to temperatures well above their glass transition temperature, T_g , they soften and exhibit a viscosity low enough to flow and consolidate the composite. In general, they must be heated to much higher temperatures than thermosets, exhibit much higher melt viscosity, and require higher pressures to form well-consolidated composite laminates. Thermoplastics offer some advantages such as reprocessability, recyclability, and, in general, higher toughness. However, thermoplastics also have several limitations that have restricted their wider acceptance as matrix materials for fiber-reinforced composites. Thermoplastics have lower solvent resistance than thermosets and require more expensive processing equipment, there are fewer commercially available thermoplastic matrix preforms available than for thermosets, and modern toughened thermosets offer similar performance to thermoplastic matrix composites. For such economic and performance reasons, thermoplastics are not widely used as thermosets for advanced composite matrix polymers. Other matrix materials are metals, ceramics, glasses, and carbon. They perform the same function in composites as the polymer matrix. These materials (with the exception of carbon) are still experimental, and their combined fraction of the composite matrix materials market is insignificant. Carbon has been used since the 1970s for exotic high-temperature ablative applications such as rocket motor nozzles. The Properties of Composites and Analysis of Composites sections of this article are general and apply to these developmental composite materials. The Processing and Applications sections, however, are concerned only with polymer matrix composites.

The matrix influences the service temperature, service environment, and fabrication process for composites. Compatibility with the reinforcement is a consideration in selecting the matrix. The matrix must adhere to the reinforcement and be capable of distributing the loads applied to the composite.

The properties of a composite can be tailored by the engineer to provide a wide range of responses, which increases their usefulness. Composites can be made to exhibit some interesting responses when loaded: They can be designed to twist and bend when loaded in plane and to extend or contract when loaded in bending. Analysis approaches are available for predicting these responses.

There are many processes for the fabrication of composites. These often result in reduction in number of parts, reduction in production time, and savings in overall manufacturing cost. The number of industries using composites and the various uses of composites continues to grow. It is difficult to foresee what the future of this class of materials will be.

II. CONSTITUENT MATERIALS

A composite can contain several chemical substances. There are additives, for example, to improve processability and serviceability. However, the two principal constituents that are always present in advanced composites are the matrix and the reinforcement. Generally, they are combined without chemical reaction and form separate and distinct phases. Ideally, the reinforcement is uniformly distributed throughout the matrix phase. The combination of the properties of the reinforcement, the form of the reinforcement, the amount of reinforcement, and matrix properties gives the composite its characteristic properties.

The matrix phase contributes to several characteristics of the composite. The matrix provides some protection for the reinforcement from deleterious environmental conditions such as harmful chemicals. The matrix plays an important role in determining the physical and thermo-physical properties of the composite. In continuous filament, unidirectionally reinforced composites, the properties transverse to the filaments are strongly influenced by the properties of the matrix. The distribution of the applied load throughout the composite is influenced by the properties of the matrix.

Table I shows typical values of selected properties of common matrix materials. The properties are tensile strength, F^{tu} , Young's modulus, E' , total strain (or strain-to-failure), ε' , coefficient of thermal expansion, α , and specific gravity. It can be seen that there is a wide variation in these values between types of matrix materials.

TABLE I Matrix Materials

Property	Epoxy	Polyimide	Polyester	Polysulfone	Polyether ether ketone	Al 2024	Ti 6-4
E^{tu} (MPa)	6.2–103	90	21–69	69	69	414	924
E^t (GPa)	2.8–3.4	2.8	3.4–5.6	2.8	3.6	72	110
ϵ^t (%)	4.5	7–9	0.5–5.0	50–100	2.0	10	8
α ($10^{-6} \text{ m m}^{-1} \text{ K}^{-1}$)	0.56	0.51	0.4–0.7	0.56	0.5	24	9.6
Specific gravity	1.20	1.43	1.1–1.4	1.24	1.2	2.77	4.43

There is great variety in polymers typically used for composite matrix materials. As discussed earlier, thermosets and thermoplastics make up the two general families of engineering polymers; but there are many different polymers within each family that exhibit very diverse properties, depending on their chemical composition. Thermosets are generally named for the characteristic reactive group of the resin (e.g., epoxy, maleimide), whereas thermoplastics are generally named for either their building block ("mer" unit; e.g., polystyrene, polyethylene, polypropylene, polyvinyl chloride) or for a characteristic repeating chemical group within the thermoplastic polymer (e.g., polysulfone, polyimide). It is more appropriate to refer to the matrix polymer as a resin system, the system being a mixture of the base polymer (or thermoset resin and curing agents). Diluents, fillers, tougheners, and other modifiers are sometimes added to the resin system to alter viscosity, increase toughness, modify reactivity of the thermosets, or change other properties of the base polymer system. Because there are so many starting combinations, it is easy to see how there can be a wide variation in the properties of materials in the same general class (e.g., based on the same basic polymer, but with different additives). The other principal constituent of a composite is the reinforcement. There are several types of materials, and their various forms are used as reinforcements. The continuous fiber has been used most extensively for the development of advanced composites. This form of reinforcement provides the highest strength and modulus. It can be used to make other forms such as woven

fabric, chopped fibers, and random fiber mats. These reinforcement forms typically reduce the mechanical performance compared to unidirectional fibers, but offer benefits in fabrication. Glass, graphite, and polymeric fibers are generally produced as bundles of many filaments of very small diameter. Metal, boron, and ceramic reinforcements are usually single fibers. After fabrication, fibers are processed with surface treatments for protection during handling and weaving and also for chemical compatibility with the matrix systems. After forming and treating, the filaments are typically wound on spools for use by manufacturers in fabricating composites, producing unidirectional preforms, or weaving into various geometries of textile preforms.

Table II lists the properties of some of the fibers, measured in the longitudinal direction (along the axis of the fiber), used in composite materials: tensile strength F^{tu} , Young's modulus E^t , coefficient of expansion α , strain-to-failure ϵ^t , diameter, and density ρ . Mechanical properties transverse to the longitudinal axis are not shown. Because of the small diameter of the fibers, transverse properties are not measured directly. Variations in the fiber properties can be caused by several factors. There can be variations in the composition of the starting material such as in E-, S-, and C-glass fibers. There can be variations in processing such as in the way the processing temperature is changed to vary the strength and modulus of graphite fibers. Also, the difficulty of performing mechanical testing on fibers contributes to uncertainty and scatter in the measured properties of fibers.

TABLE II Fiber Materials

Property	Boron	Carbon	Graphite	Aramid	Alumina	Silicon carbide	E-glass	S-glass
E^{tu} (MPa)	2.8–3.4	0.4–2.1	0.81–3.6	2.8	1.4	3.3	3.4	4.6
E^t (GPa)	379–414	241–517	34–552	124	345–379	427	69	83
α ($10^{-6} \text{ m m}^{-1} \text{ K}^{-1}$)	4.9	−0.09	−0.09	−4.0	3.4	.40	5.1	3.4
ρ (g cm^{-3})	2.5–3.3	1.55	1.55	1.60	3.90	3.07	2.55	2.5
Diameter (10^{-3} m)	0.05–0.2	0.008	0.008	0.013	0.38–0.64	0.14	0.005–0.013	0.009–0.010
ϵ^t (%)	0.67	1.0–2.0	0.4–2.0	2.5	0.4	0.6	4.8	5.4

The reinforcement is the main load-bearing phase of the composite. It provides strength and stiffness. There is a direct relationship between an increase in volume fraction of reinforcement and an increase in strength and stiffness of the composite material. This relationship depends on the assumption of compatibility with the matrix and on the existence of good bonding to the fibers.

The reinforcement and matrix are combined either before or at the time of fabrication of the composite. This depends on the fabrication process. A common practice in making continuous-fiber-reinforced laminates is to combine the constituents before fabrication into a continuous “tapelike” preform that is used much like broadgoods in that shapes are cut out of the preform and fabricated into parts. To produce this preform product, fibers are combined with resin, typically by drawing the fiber bundle through a resin or resin solution bath. Several bundles of resin-impregnated fibers are then aligned and spread into very thin layers (0.127 mm thick) on a release ply backing. The resin is usually partially cured during production of the preform to reduce its “tackiness” and improve the handleability of the preform. This tapelike preform is known as prepreg, or unidirectional tape. It is an expensive method for producing a preform, but the preform is a continuous, well-characterized, well-controlled method to combine the matrix resin and the reinforcing fiber. After prepregging, the material is usually stored in a freezer to retard the chemical reaction until the material is used. If the matrix system is a thermoplastic polymer, then no reaction can occur, and the material may be stored indefinitely at room temperature. These layers of unidirectional fibers and resin are used to make laminates by stacking many layers in directions specified by the engineer. The number of “plies” in a laminate and the direction of fibers in each layer is dependent on the mechanical properties required for the part.

The next two sections, Properties of Composites and Analysis of Composites, describe how an engineer would design a composite laminate to have the properties needed for an application. It is exactly this tailorability that makes composites attractive for engineering applications.

III. PROPERTIES OF COMPOSITES

In many of the applications in which composite materials are used, they can be considered to be constructed of several layers stacked on top of one another. These layers, or laminae, typically exhibit properties similar to those of orthotropic materials. Orthotropic materials have three mutually perpendicular planes of material property symmetry. [Figure 3](#) shows a lamina with its coordinate system and two of the planes of symmetry. We will first discuss the

properties of the lamina and some factors that influence them. Next, the properties of laminates will be discussed.

The lamina is made of one thickness of reinforcement embedded in the matrix. The elastic and strength properties of the reinforcement and the elastic and strength properties of the matrix combine to give the lamina its properties. In addition to the properties of the constituents, the amount of reinforcement, the form of the reinforcement, and the orientation and distribution of the reinforcement all influence the properties of the lamina.

The reinforcement provides the strength and stiffness of the composite. Increasing the amount of reinforcement increases the strength and stiffness of the composite in the direction parallel to the reinforcement. The effect of the form of the reinforcement is not as simple. However, some general observations can be made. Laminae reinforced by long, continuous, parallel fibers have greater strength and stiffness than laminae reinforced by short, randomly oriented fibers. Woven fiber-reinforced laminae usually have greater strength perpendicular to the principal fiber direction than do unwoven fiber-reinforced laminae. The strength and stiffness of laminae reinforced by unwoven continuous fibers decrease as the angle of loading changes from parallel to the fibers to perpendicular to the fibers.

[Table III](#) shows typical values for some properties of composite materials made of unwoven continuous fiber reinforcements. The table shows the strength and elastic properties of a laminate made of several laminae stacked on top of one another with all the fibers aligned in the same direction. The properties in the direction parallel to the fibers are much greater than the properties in the direction perpendicular to the fibers. This variation of properties with the orientation of the lamina axis is called anisotropy.

The single lamina serves as a building block. The engineer can select the orientation and number of each of the laminae in a laminate and design the laminate such that it has the required response. This designing of a laminate has some interesting implications that the engineer should understand. Two important factors are balance and symmetry.

Balance and symmetry simplify the analysis of the laminate and give it more conventional response characteristics. Balance in a laminate means that for each lamina with a positive angle of orientation there must be a lamina with an equal negative angle of orientation. Both laminae must have the same mechanical and physical characteristics. This is important in controlling the laminate's overall response to loading both in service and in fabrication. Symmetry means that for every lamina above the midplane of the laminate there is a lamina an equal distance below the midplane that is of the same type with the same orientation. Symmetry also influences the laminate response to loads.

TABLE III Typical Properties of Composite Materials: Laminates Reinforced With Unidirectional Continuous Fibers

Property	Unit	E-glass epoxy	Aramid epoxy	Graphite epoxy	Boron epoxy
Parallel to the fibers					
Tensile strength σ_x^T	MPa	1100	1380	1240	1296
Tensile modulus E_x^T	GPa	39.3	75.8	131	207
Poisson's ratio ν_{xy}	—	0.25	0.34	0.25	0.21
Total strain ε^T	%	2.2	1.8	1.21	0.66
Compressive strength σ_x^c	MPa	586	276	1100	2426
Compressive modulus E_x^c	GPa	39.3	75.8	131	221
Shear strength τ_{xy}	MPa	62.0	44.1	62.0	132
Shear modulus G_{xy}	GPa	3.45	2.07	4.83	6.2
Transverse to the fibers					
Tensile strength σ_y^T	MPa	34.5	27.6	41.4	62.7
Tensile modulus E_y^T	GPa	8.96	5.5	6.2	18.6
Compressive strength σ_y^c	MPa	138	138	138	310
Compressive modulus E_y^c	GPa	8.96	5.5	6.2	24.1
Specific gravity	—	2.08	1.38	1.52	2.01
Fiber volume V_f	%	~50	~60	~62	~50

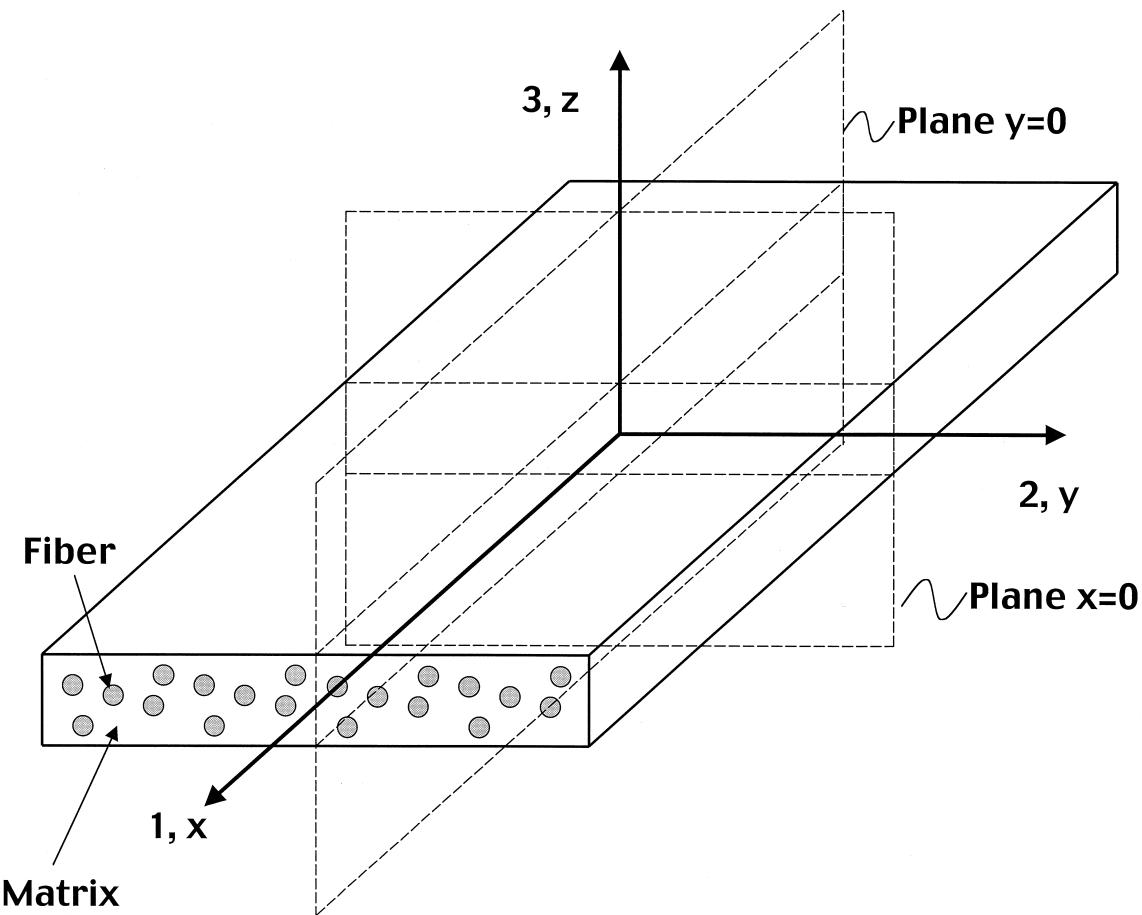


FIGURE 3 Lamina coordinate axis and planes of symmetry.

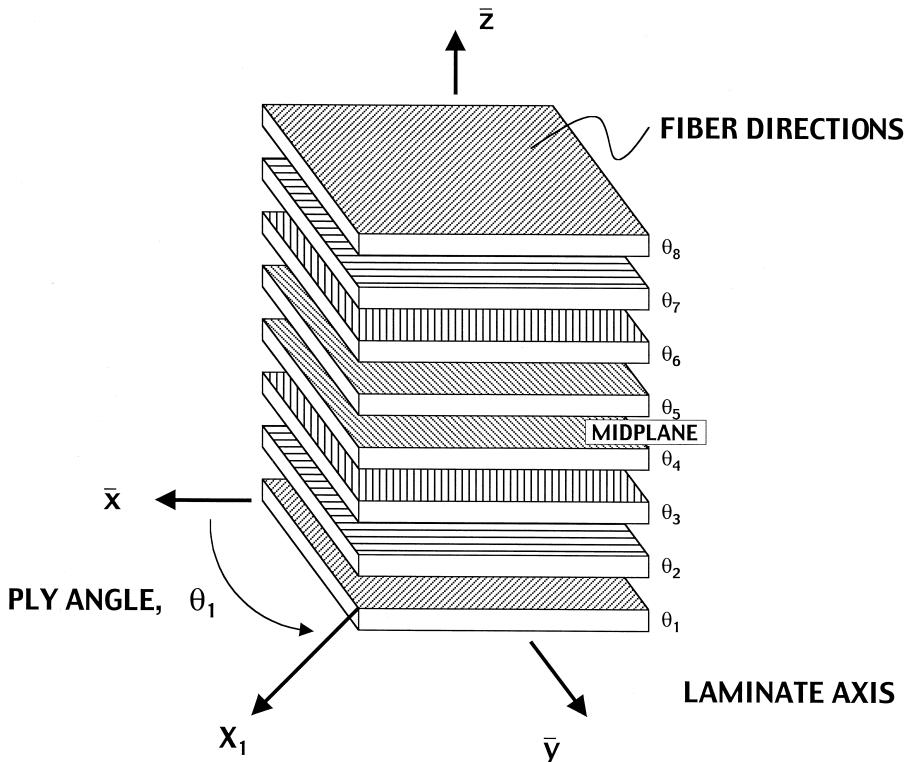


FIGURE 4 Orientation and location of laminae in a laminate.

If a laminate is not balanced and symmetrical, it will twist or bend when in-plane loads are applied. Laminates may also extend or contract when bending loads are applied. Whether the results are good or bad depends on whether they were planned or unplanned during the design of the laminate. Figure 4 shows how the laminae are oriented and stacked in a laminate.

IV. ANALYSIS OF COMPOSITES

Composite materials are complex. The properties of the constituents are different, and the fiber properties are anisotropic. The composite may also be constructed by layers, with the fiber directions varying layer to layer. Analysis of the mechanical properties of such laminates is a sophisticated process; research into better methods to predict composite performance is being pursued. However, acceptable engineering analysis methods have been developed that allow structural parts to be designed with composite materials. Further research is required to develop sound engineering methods to predict failure in composite materials, especially when subjected to severe environments that may degrade the matrix, the reinforcement, or the interfaces of the composite material. In this section, a brief summary of the currently accepted

approach to performing stress analysis of composites is presented.

The emphasis has been focused on unidirectional fiber-reinforced composites. The lamina or ply form of advanced composites has been developed into the basic unit for analysis. Most of the structural applications of advanced composites involve material in a laminated form. The laminates are constructed of plies or laminae laid up to a designed configuration (see Fig. 4).

The approach to the analysis of composites starts with the lamina and its elastic properties. Then these are related to the geometry of the lay-up for the laminate. The elastic properties and orientation of the laminae are used to calculate the modulus and stiffness of the laminate. The constitutive relationship and a selected failure criterion are used to estimate failure.

In developing the analysis of the lamina, several assumptions were made. It was assumed that (1) the fibers and matrix were bonded together, (2) the lamina was void free, (3) the lamina's thickness was small in comparison with its width and length, (4) the lamina was a homogeneous orthotropic material, and (5) the fibers were uniformly distributed within the matrix.

The lamina is analyzed as a macroscopic, homogeneous, orthotropic material in a plane stress condition. If the coordinate axes for the laminate are oriented parallel

and transverse to the fiber axis (see Fig. 3), the constitutive equation relating stress σ and strain ε is

$$\begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \tau_{12} \end{bmatrix} = \begin{bmatrix} Q_{11} & Q_{12} & 0 \\ Q_{12} & Q_{22} & 0 \\ 0 & 0 & Q_{66} \end{bmatrix} \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \gamma_{12} \end{bmatrix} \quad (1)$$

where Q is called the reduced stiffness and is defined as

$$\begin{aligned} Q_{11} &= \frac{E_1}{1 - \nu_{12}\nu_{21}}; & Q_{22} &= \frac{E_2}{1 - \nu_{12}\nu_{21}} \\ Q_{12} &= \frac{\nu_{12}E_2}{1 - \nu_{12}\nu_{21}}; & Q_{66} &= G_{12} \end{aligned} \quad (2)$$

where E_1 is Young's modulus in the direction parallel to the fibers; E_2 is Young's modulus in the direction perpendicular to the fibers; ν_{12} and ν_{21} are the major Poisson's ratio and minor Poisson's ratio, respectively; and G_{12} is the in-plane shear modulus.

Equation (1) can be inverted to give the form

$$\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \gamma_{12} \end{bmatrix} = \begin{bmatrix} S_{11} & S_{11} & 0 \\ S_{12} & S_{22} & 0 \\ 0 & 0 & S_{66} \end{bmatrix} \begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \tau_{12} \end{bmatrix} \quad (3)$$

where the S terms are the compliance coefficients and are defined as

$$\begin{aligned} S_{11} &= 1/E_1; & S_{22} &= 1/E_2 \\ S_{12} &= -\nu_{12}/E_1; & S_{66} &= 1/G_{12} \end{aligned} \quad (4)$$

Equation (4) relates the compliance coefficients to the engineering constants. These can be determined by mechanical testing. Also, estimates of the engineering constants can be made by using equations developed by micromechanics. In this approach, the properties of the constituents are used in equations for the engineering constants. These are

$$\begin{aligned} E_1 &= E_f V_f + E_m V_m \\ \nu_{12} &= \nu_f V_f + \nu_m V_m \\ P/P_m &= (1 + \xi \eta V_f)/(1 - \eta V_f) \\ \eta &= \frac{(P_f/P_m) - 1}{(P_f/P_m) + \xi} \end{aligned} \quad (5)$$

where V_f , V_m are the volume fraction of the fiber and matrix, respectively; ν_f , ν_m are Poisson's ratio of the fiber and matrix, respectively; P is the composite modulus E_2 , G_{12} , or G_{23} ; P_f is the corresponding fiber modulus E_f , G_f , or ν_f , respectively; P_m is the corresponding matrix modulus E_m , G_m , or ν_m , respectively; and ξ is a factor related to the arrangement and geometry of the reinforcement; for square packing $\xi = 2$, and for hexagonal packing $\xi = 1$.

Because not all laminae in a laminate are oriented with the fibers parallel or transverse to the laminate coordinate axis $x-y$, there must be a way to find the properties of the

lamina in the laminate coordinate systems. This is done through a transformation. By a combination of mathematical transformation and substitution, the following relationship between stress and strain for an arbitrary lamina k is developed:

$$\begin{bmatrix} \sigma_x \\ \sigma_y \\ \tau_{xy} \end{bmatrix}_k = \begin{bmatrix} \bar{Q}_{11} & \bar{Q}_{12} & \bar{Q}_{16} \\ \bar{Q}_{12} & \bar{Q}_{22} & \bar{Q}_{26} \\ \bar{Q}_{16} & \bar{Q}_{26} & \bar{Q}_{66} \end{bmatrix}_k \begin{bmatrix} \varepsilon_x \\ \varepsilon_y \\ \gamma_{xy} \end{bmatrix}_k \quad (6)$$

The \bar{Q} terms are the components of the stiffness matrix for the lamina referred to an arbitrary axis. They are defined as

$$\begin{aligned} \bar{Q}_{11} &= Q_{11} \cos^4 \theta + 2(Q_{12} + 2Q_{66}) \sin^2 \theta \cos^2 \theta \\ &\quad + Q_{22} \sin^4 \theta \\ \bar{Q}_{22} &= Q_{11} \sin^4 \theta + 2(Q_{12} + 2Q_{66}) \sin^2 \theta \cos^2 \theta \\ &\quad + Q_{22} \cos^4 \theta \\ \bar{Q}_{12} &= (Q_{11} + Q_{22} - 4Q_{66}) \sin^2 \theta \cos^2 \theta \\ &\quad + Q_{22} (\sin^4 \theta + \cos^4 \theta) \\ \bar{Q}_{66} &= (Q_{11} + Q_{22} - 2Q_{12} - 2Q_{66}) \sin^2 \theta \cos^2 \theta \\ &\quad + Q_{66} (\sin^4 \theta + \cos^4 \theta) \\ \bar{Q}_{16} &= (Q_{11} - Q_{12} - 2Q_{66}) \sin^2 \theta \cos^3 \theta \\ &\quad + (Q_{12} - Q_{22} + 2Q_{66}) \sin^3 \theta \cos \theta \\ \bar{Q}_{26} &= (Q_{11} - Q_{12} - 2Q_{66}) \sin^2 \theta \cos \theta \\ &\quad + (Q_{12} - Q_{22} + 2Q_{66}) \sin \theta \cos^3 \theta \end{aligned} \quad (7)$$

where θ is the ply angle according to the Tsai convention (see Fig. 4). Counterclockwise rotations are positive and clockwise rotations are negative.

The constitutive relationships for the lamina and linear small deformation theory were used to develop the analysis for composite structures. Some assumptions that were made are as follows: (1) The laminae are bonded together, and they do not slip relative to one another when load is applied; (2) the normals to the undeformed midplane of the laminate are straight, and they remain so with no change in length after deformation; (3) the thickness of the plate is small compared with the length and width; and (4) the strain in the thickness direction is negligible. The in-plane strain is assumed constant for all the laminae. The stress varies from lamina to lamina. As a simplification, the force and moment resultants were defined. The force resultants N_x , N_y , and N_{xy} were defined as the sum of the laminae stresses per unit width. The moment resultants M_x , M_y , and M_{xy} were defined as the sum of the respective stresses, times the area over which they act, multiplied by the appropriate moment arm. The in-plane strains at the

midplane, ε_x^0 , ε_y^0 , and γ_{xy}^0 , and the curvatures, κ_x , κ_y , and κ_{xy} , are related to the resultants as shown in Eq. (8).

$$\begin{bmatrix} N_x \\ N_y \\ N_{xy} \\ \hline M_x \\ M_y \\ M_{xy} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & A_{16} & | & B_{11} & B_{12} & B_{16} \\ A_{12} & A_{22} & A_{26} & | & B_{12} & B_{22} & B_{26} \\ A_{16} & A_{26} & A_{66} & | & B_{16} & B_{26} & B_{66} \\ \hline B_{11} & B_{12} & B_{16} & | & D_{11} & D_{12} & D_{16} \\ B_{12} & B_{22} & B_{26} & | & D_{12} & D_{22} & D_{26} \\ B_{16} & B_{26} & B_{66} & | & D_{16} & D_{26} & D_{66} \end{bmatrix} \begin{bmatrix} \varepsilon_x^0 \\ \varepsilon_y^0 \\ \gamma_{xy}^0 \\ \hline \kappa_x \\ \kappa_y \\ \kappa_{xy} \end{bmatrix} \quad (8)$$

where N_x , N_y , and N_{xy} are force resultants; M_x , M_y , and M_{xy} are moment resultants; $[A]$ is the in-plane stiffness matrix for a laminate; $[B]$ is the coupling stiffness matrix for a laminate; $[D]$ is the bending stiffness matrix for a laminate; ε_x^0 , ε_y^0 , and γ_{xy}^0 are the strains at the laminate geometric mid-plane; and κ_x , κ_y , and κ_{xy} are the curvatures of the laminate.

Examination of Eq. (8) shows that the $[A]$ matrix is the coefficients for the in-plane strains. The $[B]$ matrix relates the curvatures to the force resultants and the in-plane strains to the moment resultants. The $[D]$ matrix relates the curvatures to the moment resultants. Equation (8) can be partially or fully inverted, depending on whether the

strains, curvatures, forces, or moments are known in a given situation.

The definitions for the elements of the $[A]$, $[B]$, and $[D]$ matrices are

$$A_{ij} = \sum_{k=1}^n (\bar{Q}_{ij})_k (h_k - h_{k-1}) \quad (9)$$

$$B_{ij} = \frac{1}{2} \sum_{k=1}^n (\bar{Q}_{ij})_k (h_k^2 - h_{k-1}^2) \quad (10)$$

$$D_{ij} = \frac{1}{3} \sum_{k=1}^n (\bar{Q}_{ij})_k (h_k^3 - h_{k-1}^3) \quad (11)$$

[Figure 5](#) shows how k and h are defined for the laminae.

The force resultants and moment resultants are defined as

$$\begin{bmatrix} N_x \\ N_y \\ N_{xy} \end{bmatrix} = \int_{-h/2}^{h/2} \begin{bmatrix} \sigma_x \\ \sigma_y \\ \tau_{xy} \end{bmatrix} dz \quad (12)$$

and

$$\begin{bmatrix} M_x \\ M_y \\ M_{xy} \end{bmatrix} = \int_{-h/2}^{h/2} \begin{bmatrix} \sigma_x \\ \sigma_y \\ \tau_{xy} \end{bmatrix} z \, dz \quad (13)$$

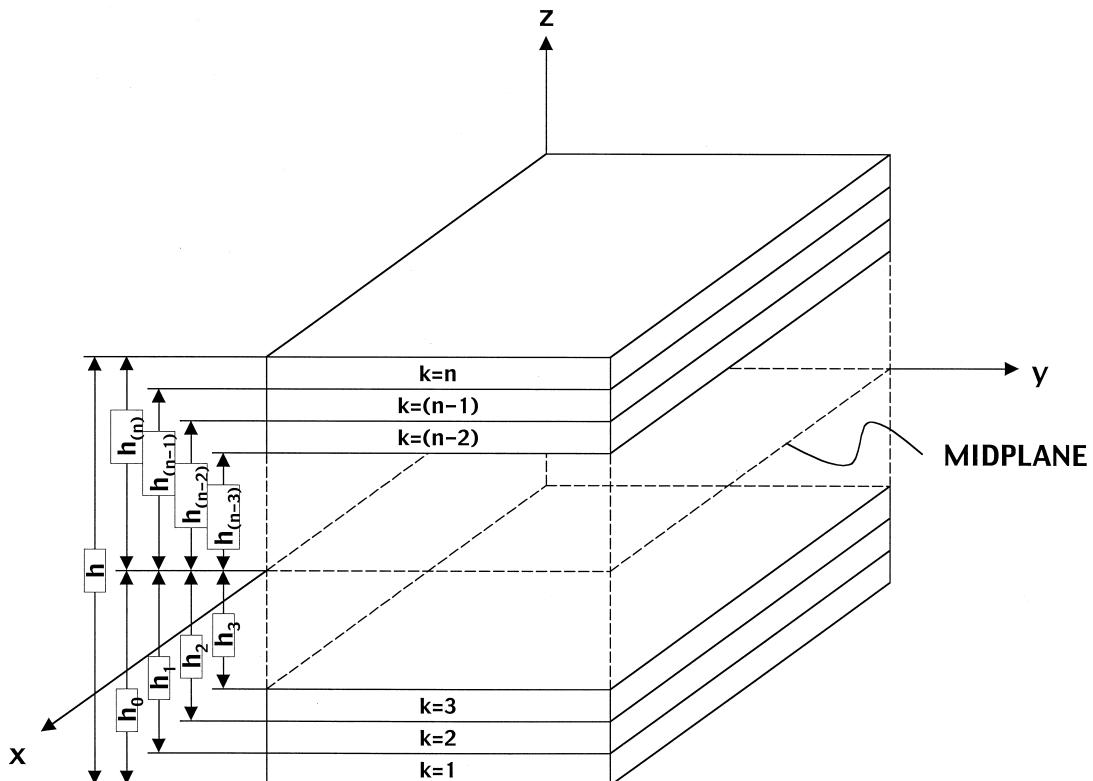


FIGURE 5 Relationship of laminae to the laminate coordinates.

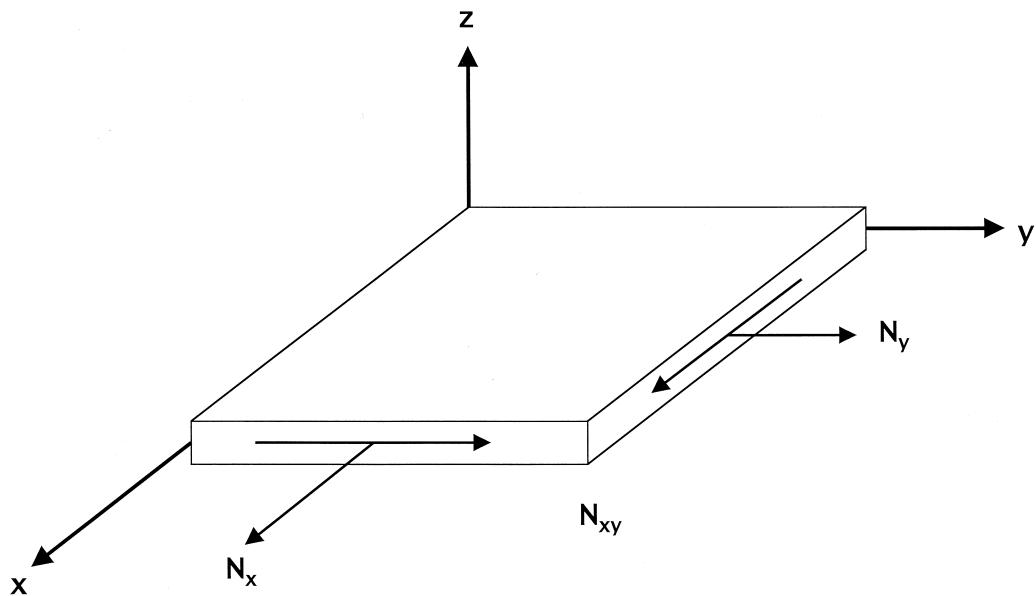


FIGURE 6 Force resultants on an element.

where σ_x , σ_y , and τ_{xy} are the stresses in the laminate coordinate system and z is the distance from the midplane in the direction normal to the midplane. Figures 6 and 7 show how the force and moment resultants act on an element in the laminate.

Equation (8) is the constitutive equation for a general laminated plate. Significant simplifications of Eq. (8) are possible. If the $[B]$ is made zero, the set of equations for the stress and moment resultants is uncoupled. “Uncou-

pled” means that in-plane loads generate only in-plane responses, and bending loads generate only bending responses. The $[B]$ can be made zero if for each lamina above the midplane there is a lamina with the same properties, orientation, and thickness located at the same distance below the midplane. This is significant not only in simplifying the calculations but also in the physical response to load and in fabrication. If the $[B]$ is zero, the laminate will not warp when cured, and no bending will be induced

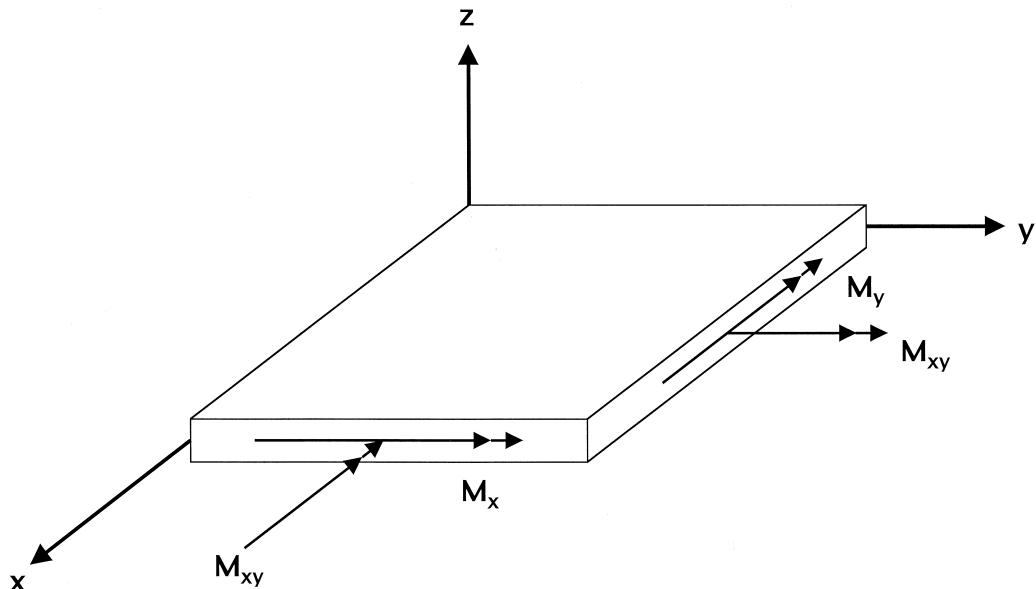


FIGURE 7 Moment resultants on an element (following the right-hand rule).

when the laminate is under in-plane loads. Equation (8) becomes

$$\begin{bmatrix} N_x \\ N_y \\ N_{xy} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & A_{16} \\ A_{12} & A_{22} & A_{26} \\ A_{16} & A_{26} & A_{66} \end{bmatrix} \begin{bmatrix} \varepsilon_x^0 \\ \varepsilon_y^0 \\ \gamma_{xy}^0 \end{bmatrix} \quad (14)$$

and

$$\begin{bmatrix} M_x \\ M_y \\ M_{xy} \end{bmatrix} = \begin{bmatrix} D_{11} & D_{12} & D_{16} \\ D_{12} & D_{22} & D_{26} \\ D_{16} & D_{26} & D_{66} \end{bmatrix} \begin{bmatrix} k_x^0 \\ k_y^0 \\ k_{xy}^0 \end{bmatrix} \quad (15)$$

In the preceding discussion, only the elastic properties of the laminate were considered. The elastic behavior of a laminate can be used to analyze the strength behavior of a laminate. To determine the strength of a laminate, we need a failure criterion for the lamina. It is assumed that the response of the lamina will be the same when it is in the laminate under the same stresses or strains. The strength of the laminate will be related to the strength of the individual lamina. The general approach is to determine the force and moment resultants or the mid-plane strains and curvatures for the laminate by using the laminate plate equation or an inverted form. The stress or strain is calculated for each lamina in the laminate axis system, and then it is transformed into the lamina axis system for each lamina and the failure criteria applied to determine if failure occurred in the lamina. If the first-ply failure concept for the laminates is applied, the laminate is considered to have failed when the first lamina fails. No single approach has been universally accepted for strength analysis of laminates after first-ply failure.

V. FABRICATION OF COMPOSITES

Fabrication of components from composite materials is somewhat different from that using traditional engineering materials in that the properties of a composite are highly dependent on the geometry of the reinforcement. The structural designer must consider the issues associated with processing the composite part to ensure that reinforcement volume fraction, reinforcement geometry, and other material properties can be produced economically. The diversity of composite applications has stimulated the development of a wide range of techniques for fabricating structural composites. In fact, one of the principal reasons for the success of composites is the ease of fabrication and the many different processes with widely varying levels of sophistication and cost that are available for their production. Structural and decorative composites can be fabricated with techniques ranging from very crude hand lay-up processes without molds to very

sophisticated techniques with complex molds, woven 3D reinforcement preforms, and artificial intelligence-guided computer-controlled resin infusion and curing. The configuration of the part, along with the basic manufacturing considerations such as volume, production speed, and market conditions, determine whether a part will be built in open or closed molds, by compression molding, or by an automated system. Composite fabrication technologies can be classified as either open or closed molding, the choice of appropriate technique being governed by factors mentioned earlier.

We can group most of the processes into two classes: open molding and closed molding. The main distinction is that open molds are one piece and use low pressure or no pressure, and closed molds are two pieces and can be used with higher pressure.

A. Open-Mold Processes

Open-mold processes such as spray-up, wet hand lay-up, autoclave, filament winding, vacuum infusion, pultrusion, or combinations of these techniques are the most common open-mold methods to produce composite products. Many products are suited to these manufacturing methods, including aerospace structures, tanks, piping, boat hulls and structures, recreational vehicle components, commercial truck cabs and components, structural members, and plumbing applications (e.g., tubs, showers, pools, and spas).

In spray-up and wet hand lay-up open molding, the mold surface typically has a high-quality finish and is the visual surface of the finished part. Common to all open molding techniques is mold preparation. To prepare the mold surface prior to spray-up, hand lay-up, or vacuum infusion, the mold is treated with a release agent to aid in composite part removal and then may be coated with a “gel coat” (a color-tinted layer of resin that becomes the visual surface of the finished part).

In spray-up fabrication, the thermoset resin is sprayed into the prepared mold simultaneously with chopped reinforcing fiber. The random sprayed-up mat of fiber and resin may then be compacted with hand rollers prior to cure to produce a more dense part. A hand lay-up component, the resin, and reinforcement (usually a fabric or random fiber mat) are laid into the mold, compacted with rollers, and allowed to cure. Often hand lay-up is combined with spray-up techniques depending on the structural requirements of the part. Fiber volumes of 15 to 25% are typically achieved with these techniques. There are several variations of the basic process. A vacuum bag made of a nonporous, nonadhering material can be placed over the lay-up. Then a vacuum is drawn inside the bag. The atmospheric pressure outside the bag eliminates the voids

and forces out entrapped air and excess resin. Another approach is to use a pressure bag. The bag is placed against the lay-up and the mold covered with a pressure plate. Air or steam pressure is applied between the bag and the plate.

Vacuum infusion is an open molding process that is very suitable for large components for many important reasons. Vacuum infusion uses an airtight membrane over the entire part to provide vacuum pressure on the reinforcement and to prevent any volatile resin products from escaping into the atmosphere. The resin is introduced after the entire reinforcement is laid into the mold and the vacuum membrane is in place; this reduces some issues associated with the working time of the resin prior to cure. Finally, higher volume fractions of reinforcement are achievable since the reinforcement is compacted by vacuum pressure and only the minimum amount of resin necessary is added. Reinforcement volume fractions up to 70% have been reported.

An open-mold technique that is widely used in the aerospace industry and is slightly different from the preceding processes is autoclaving. One difference in this process is that the entire assembly (the lay-up and supporting unit) is placed inside an autoclave. An autoclave is a large pressure vessel that is used to provide heat and pressure to the lay-up during cure. Autoclaves are usually cylindrical, with an end that opens for full access to the interior. They have provision to pull vacuum on the lay-up assembly, and they often have multiple temperature sensors that are used to monitor the temperature of the part during cure. The curing takes place under pressure, 1–10 bar (15–150 psi), and at elevated temperature. The lay-up assembly is slightly different (Fig. 8). The top surface of the lay-up is covered with a perforated or porous release film, and if necessary bleeder plies of dry cloth are added to absorb excess resin. Then the assembly is sealed within a nonporous sheet material and placed into the autoclave. The application of pressure and control of temperature is critical. This process offers better quality control than other low- or no-pressure molding processes.

Another process that is used extensively is filament winding. The concept of wrapping filaments around articles to improve their performance is very old. The modern practice of filament winding was developed in response to the requirements for lightweight pressure vessels. Filament winding uses continuous reinforcement to maximize the use of fiber strength. Preimpregnated tape, or a single strand that has passed through a resin bath, is wound onto a mandrel in a prescribed pattern. Design and winding technique allow the maximum fiber strength to be developed in the direction desired. When the winding is completed, the assembly is cured either at room temperature or in an oven. After cure, the mandrel is removed. This process provides for a high level of fiber content.

The process of pultrusion is the opposite of extrusion. The reinforcement is passed through a resin bath and then pulled through a die that controls the resin content and final shape. The die can be heated to cure the resin, or the material can be passed through an oven for curing.

B. Closed-Mold Processes

The closed-mold processes use a two-part mold or die. When the two parts are put together, they form a cavity in the shape of the article to be molded. The molds are usually made of metal with smooth cavity surfaces. Higher pressures and temperatures than those in open molding are usually used. The processes produce very accurate moldings. Most of the processes are attractive for mass production.

Matched die molding is a closed-mold process. There are variations to this process. The main variations concern the form of the starting material and the manner in which it is introduced into the mold. In some cases, the reinforcement is first made into a preform and placed into the mold and then a metered amount of resin is added—this is known as resin transfer molding, or RTM. RTM is a widely used technique for production of components that require accurate dimensional tolerances, since the outer

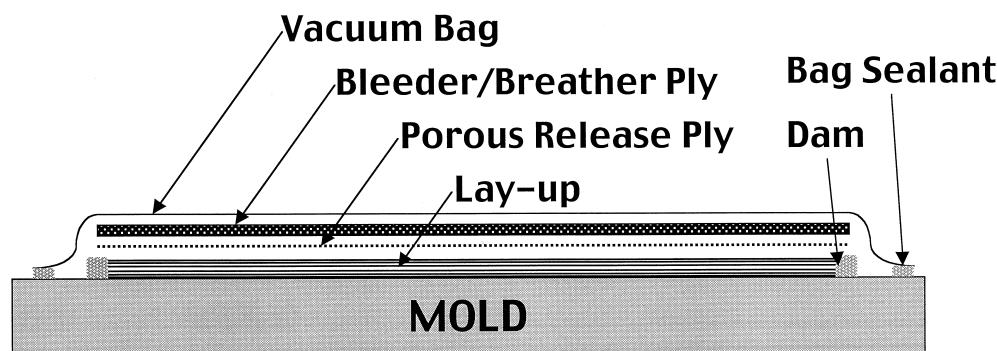


FIGURE 8 Cross section of the composite laminate lay-up and vacuum bagging processing method.

surface of the part is determined by the tool surface. In other cases, a resin-reinforcement mixture is made and a premeasured amount placed into the mold. The molding compound can be introduced automatically or manually. The molding temperatures range from 100°C (212°F) to 140°C (284°F). Pressures range from 7 to 20 bar. Cure cycles can be as short as minutes.

The selection of a fabrication process depends on several factors, including the materials to be processed, the size and design of the article, the number of articles, and the rate of production. Processes differ in their capacity to use different forms of reinforcement and to achieve the proper distribution and amount of reinforcement. The chemistry and rheology of the resin are important factors in process selection. Closed molds require higher temperatures and pressures.

The size and shape of the article to be produced affect the selection. Very large articles such as boat hulls and vehicle bodies and components are more easily and economically produced in open-mold processes. Small gears and precision electrical parts are more suitably produced in closed molds. Shapes that are surfaces of revolution are ideal for filament winding. Very large cylindrical containers have been fabricated by this process. In most open-mold processes, the molds are made of low-cost materials and are easily fabricated but have shorter lives. Autoclave processing of composites, while considered an open-mold technique, requires accurate, robust tools because of the relatively high temperatures and pressures used in the autoclave. Autoclave techniques are well suited to large structural components for aerospace applications; hence, dimensional accuracy of the tools is critical. Open-mold, hand lay-up processes have higher labor cost. If one is making a large number of parts and requires high production rates, mold life and labor cost are important factors. Open-mold processes are usually more costly in these two areas than closed-mold processes. Also, some closed-mold processes can be automated.

Automating the fabrication of advanced composites and improving processing science for composites are two current goals. The advantages of advanced composites are lighter weight, higher strength- and modulus-to-weight ratios, flexibility in design and fabrication, and usually fewer parts per component. Automating the fabrication process could result in a reduction in labor cost and an improvement in quality. The computer-aided manufacturing technology could be utilized to reduce the total labor hours. The application of higher precision control technology could improve quality and lower rejection rates. Work in processing science should result in increased understanding of the cure process, which will aid the development of resin systems and automating production cycles.

Fabrication processes for other matrix materials are important for the use and continued development of these composites. However, not as much work has been done in these areas. The use of these materials represents a small part of the overall uses of composite materials.

VI. USES OF COMPOSITES

Composite materials have been introduced into almost every industry in some form or fashion. We shall look at some of the advantages of using composites and discuss some of the industries that have made use of these materials.

The wide range of property values attained with composites and the ability to tailor the properties is an advantage. Composite materials also generally have higher strength- and modulus-to-weight ratios than traditional engineering materials. These features can reduce the weight of a system by as much as 20 to 30%. The weight savings translates into energy savings or increased performance. Advanced composites exhibit desirable dynamic properties and have high creep resistance and good dampening characteristics. In fact, the superior fatigue performance of composite materials enables them to be used to repair metallic airframes with fatigue damage.

Since composite materials can be manufactured into almost any shape, they allow great design flexibility and offer reduced parts count for articles. The opportunity to select the constituents, tailor them to obtain the required properties, and then through design make the optimum use of the properties is a situation that makes composites very attractive to many industries.

The matrix polymer can impart great chemical and corrosion resistance to composites. The transportation industry has made extensive use of composite materials. The light weight and high strength and the ability to easily manufacture aerodynamic shapes have resulted in lower fuel costs. The lack of corrosion of the materials and the low maintenance cost have reduced the cost of ownership and extended the service life of many parts and products. Examples of products in this industry include auto and truck bodies and parts, trailers, tanks, special-purpose vehicles, and manufacturing equipment.

Composites have added new dimensions to the design and construction of buildings. Their ease of manufacture, light weight, high strength, low maintenance, decorative-ness, and functionality have had a significant impact on the industry. New-construction time has been reduced and more flexibility has been added to the design of structures.

Composite materials affected the marine industry very early in their development, and their influence continues to grow. Lack of corrosion, low maintenance, and design flexibility have contributed to the acceptance of

composites. The ease of fabricating very large and strong articles in one piece has been another. In addition to pleasure boats, large military and commercial boats and ship hulls have been fabricated. Large tanks for fuel, water, and cargo have been used aboard ships. Composites have made the greatest impact in the sporting goods industry, replacing traditional materials at a revolutionary pace. Applications such as golf club shafts, fishing poles, tennis rackets, skiing equipment, boating applications, and many other sports equipment products are now produced almost exclusively using advanced composites. In most cases, the change in material has translated into an improvement in performance or safety for participants.

The aerospace and military markets are the two areas that have accounted for the largest effort in the development and advancement in composite technology. The need for stronger, stiffer, and lighter structures was an opportunity for composite materials to demonstrate their superiority over more commonly used materials. Durability and low maintenance are additional assets. These increase the service life and reduce the cost of maintaining systems. The development of new and the improvement of existing fabrication processes have brought about a reduction in manufacturing cost. There have been reductions in the number of parts required to construct some components by using molding and composite materials. The unique features of composites have enabled designers to formulate advanced systems that could be made only of composite materials. New military aircraft almost exclusively utilize advanced composites for structure. Rocket motor cases, nozzles, and nose cones are missile applications. Radar domes, rotor blades, propellers, and many secondary structure components such as fairings, doors, and access panels are also fabricated from advanced composites. Numerous pressure vessels, armaments, and items of space hardware are made of selected composite materials.

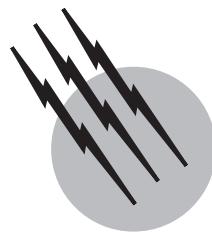
The use of composite materials will continue to grow. As more engineers come to understand composites, more opportunities will be recognized for their use. As the use of composites increases, more developments will take place in the areas of constituent materials, analysis, design, and fabrication. Composite materials offer tremendous for tailorability, design flexibility, and low-cost processing with low environment impact. These attributes create a very bright future composite materials.

SEE ALSO THE FOLLOWING ARTICLES

ADHESION AND ADHESIVES • BIOPOLYMERS • CARBON FIBERS • FRACTURE AND FATIGUE • METAL MATRIX COMPOSITES • POLYMERS, MECHANICAL BEHAVIOR • POLYMERS, THERMALLY STABLE • SANDWICH COMPOSITES

BIBLIOGRAPHY

- Ashton, J. E., Halpin, J. C., and Petit, P. H. (1969). "Primer on Composite Materials: Analysis," Technomic Publishing Company, Stamford, CT.
- Hull, D. (1981). "An Introduction to Compositional Materials," Cambridge University Press, London.
- Jones, R. M. (1975). "Mechanics of Composite Materials," Scripta Book Company, Washington, D.C.
- Sih, G. C., and Hsu, S. E. (1987). "Advanced Composite Materials and Structures," VNU Science Press, Utrecht, The Netherlands.
- Tsai, S. W. (1985). "Composites Design—1985," Think Composites, Dayton, OH.
- Tsai, S. W., and Hahn, H. T. (1980). "Introduction to Composite Materials," Technomic Publishing Company, Westport, CT.
- Whitney, J. M., Daniel, I. M., and Pipes, R. B. (1982). "Experimental Mechanics of Fiber Reinforced Composite Materials," Society for Experimental Stress Analysis, Brookfield Center, CT.
- Industry Overview: FRP Materials, Manufacturing Methods and Markets, (1999). *Composites Technol.* **5**, 6–20.



Glass

Solomon Musikant

General Electric Company

- I. History
- II. Definition
- III. Structure
- IV. Composition
- V. Annealing
- VI. Properties
- VII. Manufacture
- VIII. Applications
- IX. The Future of Glass

GLOSSARY

Annealing Thermal process whereby the residual strains in a glass are removed by a precisely controlled thermal treatment.

Batching In the making of glass, the step in the process where the initial measurement and mixing of the ingredients takes place.

Birefringence Quality of an optical material that is characterized by different indices of refraction in various directions.

Devitrification Process whereby a glassy material is converted to a crystalline material.

Dielectric constant Material property determined by the ratio of the charge-holding capacity of a pair of parallel, electrically conductive plates with the material filling the space between the plates to the charge-holding capacity of the same plates when a vacuum replaces the intervening material.

Dispersion Variation of the index of refraction in a material as a function of wavelength of the light transiting

through the medium.

Forming Process whereby a glass at a low enough viscosity to permit relatively easy deformation is shaped by application of forces into a desired configuration and then cooled to a rigid state.

Fused quartz Glassy product resulting from the melting of crystalline quartz and then cooling at a rate fast enough to avoid devitrification.

Glass-ceramic Body composed of a glassy matrix within which fine precipitates of a crystalline phase have been formed by appropriate heat treatments. Properties of the glass-ceramic are thereby substantially different than those of the initial glass.

Glass-fiber wave guide Glass fiber with its index of refraction varying radially, so that the index is higher at the axis than at the periphery and thereby able to confine a coherent ray of light, which is modulated to convey information.

Index of refraction Ratio of the velocity of light in vacuum to the velocity of light in the medium.

Network formers Those elements in the glass

composition that, during the cooling process from the molten state, are able to induce molecular structures such as polyhedra and triangles, which lead to the glassy state.

Plane-polarized light Light beam in which all the electromagnetic waves are aligned so that the electric vectors are all parallel to each other and all the magnetic vectors are parallel to each other.

Quartz Crystalline form of silica (SiO_2) that occurs in two forms: low quartz, below 573°C, and high quartz, between 573 and 867°C. High quartz can be viewed as composed of connected chains of silica tetrahedra.

Silica Term given to the stable oxide of silicon, SiO_2 , which can occur in glassy or in a wide variety of crystalline forms.

Spectral transmittance Property of a medium that permits light to pass through it without absorption or scattering.

A GLASS is an inorganic substance in a condition that is continuous with, and analogous to, the liquid state of that substance, but which, as a result of having been cooled from a fused condition, has attained so high a degree of viscosity as to be for all practical purposes rigid. The limitless possible compositions and variety of equilibrium and metastable conditions provide the keys to the widest possible range of optical, physical, and mechanical properties.

I. HISTORY

The art of glass is among the oldest of mankind's achievements preceding written history. Glass tektites of extraterrestrial origin were probably found by ancient humanity. Primitive people shaped objects from volcanic glass found in nature. Obsidian is one of the most common of such materials and is translucent and usually blackish in color, although sometimes found in red, brown, or green hues, depending on the specific composition. Its earliest use was for cutting or piercing implements or weapons, as the shards are extremely sharp and often elongated. More advanced cultures used obsidian for jewelry, ornaments, and sacred objects. Worked objects have been found in Europe, Asia Minor, and North and Central America. The flake knife, in which the glass is fashioned by cleavage to the form of a sharp blade, was used well into the age of iron. The history of glass has been reviewed by Morey.

The discovery of glass making was not recorded but probably was by accident. Fire-making knowledge provided access to high temperatures and to both metallurgy

and glass making. A possible mode of discovery was the accidental ignition of grain and the fusion (melting) of its ash. Pliny describes a Phoenician story regarding the accidental discovery of glass making when blocks of nitre were used to support a wood fire on a sand beach. The nitre (potassium nitrate) formed a low-melting eutectic with the pure silica (SiO_2) sand of the beach, fused, and upon cooling formed glass.

One of the oldest pieces of glass is a molded amulet of deep lapis lazuli color of about 7000 B.C. found in Egypt. Many early examples of glass objects have been found in Syria and in the Euphrates region that can be dated back to 2500 B.C. Mesopotamia also is a region where early glass objects have been discovered and believed to have been fashioned as early as 2700 B.C. Explorations at Ninevah have revealed Assyrian styles that give detailed descriptions of glass-making furnaces and methods of preparing glass representing the state of the art at about 650 B.C.

There was a stable glass industry in Egypt at the beginning of the eighteenth dynasty (approximately 1550 B.C.). From that time to the beginning of the Christian era, Egypt was the greatest center of glass manufacture. The industry was centered in Alexandria.

Initially, glass was considered to be a gem. As late as Ptolemaic times, glass gems were highly valued. Glass was later used for hollow vessels, such as unguent jars. The idea of pressing glass in open molds is thought to have originated about 1200 B.C.

The invention of glass blowing was a revolutionary advance. This technique is thought to have been discovered about the beginning of the Christian era. This method was accompanied by dramatic improvements in glass quality. The highest-quality glassware of the early Roman empire was characterized by transparency and freedom from color. Drinking vessels of glass superceded those of silver and gold.

The ensuing rapid expansion of glass manufacture occurred because of the combination of glass blowing and the stability afforded by the Roman Empire. Glass became a common home material. From the fall of Rome (476 A.D.) to the eleventh century, glass manufacture declined greatly. However, in the tenth century, a limited treatise on glass manufacture was written, and a far more comprehensive writing was composed in the twelfth century by Theophilus Presbyter, reflecting the influence of the Venetian renaissance.

Venice became the hub of a dominant glass industry in the eleventh century and maintained that status for four centuries. The Venetians took great pains to maintain their proprietary knowledge. From Venice, the knowledge finally spread widely. The book "L'Arte Vetraria" by Neri was published in Florence in 1612. It was later annotated

by Merrett and Kunckel and is an important source document for glass history.

Neri's book initiated the modern scientific approach to glass making by providing a basis in facts and observations. Neri was a priest, and his interests are presumed to have been academic in nature. Merrett was an English physician, a member of the royal society, and able to bring the scientist's viewpoint to the work, while Kunckel was descended from a family of glass makers. He added the practical knowledge to the work.

After 1600, glass technology and manufacturing spread rapidly throughout Europe. During this early period, the art of cut glass developed. Coal replaced wood as a fuel in 1615, and flint glass (high refractive index) was invented in 1675. The name *flint* comes from the early use of high-purity silica in the form of flint to make this glass. Lead oxide was the essential ingredient in developing the high index of refraction.

During the nineteenth century, the demands of optical instruments forced great advances in the theoretical understanding of glass properties, especially the optical properties. Optical glass must be free from imperfections, homogeneous, and available in a wide variety of optical properties that are repeatable from batch to batch.

Michael Faraday in England conducted extensive research on the properties of glass. Later, O. Schott in Germany performed an outstanding study on the effect of composition on optical properties. E. Abbe, a physicist, provided the optical property requirements that Schott attempted to achieve by new compositions of glass, while Winkelmann conducted critical studies of the mechanical properties of glass.

These brilliant researchers ushered in the modern age of glass making. The military requirements for optical glass during World War I created a new impetus to both the science and manufacturing technologies of glass. To the present day, a seemingly endless series of technology advances, inventions, and broad applications continues to occur.

II. DEFINITION

No wholly satisfactory definition of glass is possible. A reason for this difficulty is that words such as *solid* and *liquid* are generally used much more loosely than allowed by strict scientific definition. The definition of glass is often given in terms of undercooled liquid or amorphous solid. However, liquid and solid are terms applied to a substance that undergoes a phase change. In the case of glass, no phase change occurs as it cools from the fluid state, but there is a complete continuum of properties such as its viscosity. In fact, the viscosity of glass is generally

used to define its various conditions such as softening point and melting point.

Even at room temperature, although glass is "solid" in the ordinary sense of the word, in actuality it is still fluid, although with an extremely high viscosity. Cases of centuries-old windows that show signs of flow are well known. Morey provides the following definition:

A glass is an inorganic substance in a condition which is continuous with, and analogous to, the liquid state of that substance, but which, as a result of having been cooled from a fused condition, has attained so high a degree of viscosity as to be for all practical purposes rigid.

The American Society for Testing Materials (Committee C-14 on Glass and Glass Products) has adopted the following definition:

Glass—An inorganic product of fusion which has cooled to a rigid condition without crystallizing.

(a.) Glass is typically hard and brittle and has a conchoidal fracture. It may be colorless or colored, and transparent to opaque. Masses or bodies of glass may be colored, translucent, or opaque by the presence of dissolved, amorphous, or crystalline material.

(b.) When a specific kind of glass is indicated, such descriptive terms as flint glass, barium glass, and window glass should be used following the basic definition, but the qualifying term is to be used as understood by trade custom.

(c.) Objects made of glass are loosely or popularly referred to as glass; such as glass for a tumbler, a barometer, a window, a magnifier or a mirror.

III. STRUCTURE

The understanding of the structure of glass has been made possible by X-ray diffraction analysis. It was learned through such early investigations that in crystalline silica, SiO_2 , a tetrahedral relationship exists between the Si^{+4} and the O^{-2} ions. These ions form repeating tetrahedra in space where the Si^{+4} is at the center and the O^{-2} ions are located at each of the four apexes of each tetrahedron. This arrangement provides charge balance, since each O^{-2} is shared by two Si^{+4} ions. Such an arrangement is indicated in Fig. 1 and is termed fourfold coordination. In this case, the distance between the ions of opposite charge is about 1.60 Å.

Silica glass seems to be described reasonably well as a random network of Si–O tetrahedra with some variability in Si–O–Si bond angles.

A two-dimensional representation of a crystal with an A2O3 arrangement of ions is shown in Fig. 2(a). The term A represents a cation, while O is used in its normal sense to represent an oxygen anion. Figure 2(b) shows a schematic

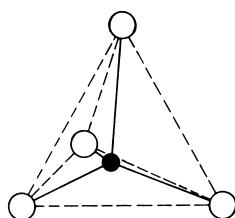


FIGURE 1 The silica tetrahedron: ● silicon ion, O oxygen ion. [From Morey, G. W. (1954). "The Properties of Glass," Reinhold, New York.]

representation of an A2O₃ glass where the A2O₃ coordination is maintained but the periodicity is not.

A. Formation

Cations such as Si, which readily form polyhedra and triangles with oxygen, are termed network formers because such coordination polyhedra lead to glass structure. Alkali silicates very readily form glasses, as schematically indicated in Fig. 3. The alkali metals are termed network modifiers, because these ions take up random positions in the network and thus "modify" or change the structure of the network. Other cations with higher valence than the alkali metals and lower coordination numbers modify the network structures less drastically and are called intermediates. Typical network formers, modifiers, and intermediates are listed in Table I, along with the valence and coordination number for each of the ions.

As alluded to earlier, X-ray diffraction work has established that the structure of silica glass is well described as a random network of SiO₄ tetrahedra with variable bond angles in the silicon–oxygen–silicon covalent connections.

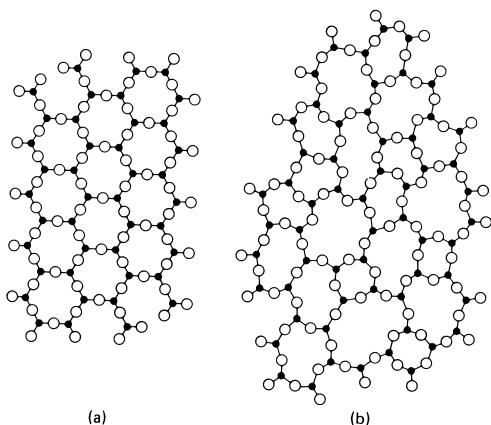


FIGURE 2 Schematic representation of (a) ordered crystalline form and (b) random-network glassy form of the same composition. [From Kingery, W. D., Bowen, H. K., and Uhlmann, D. R. (1976). "Introduction to Ceramics," Wiley, New York.]

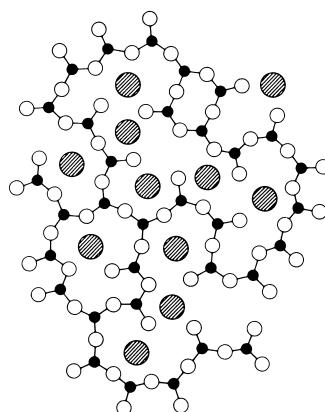


FIGURE 3 Schematic representation of the structure of a sodium silicate glass: ● Si⁴⁺, O O⁻², ○ Na⁺. [From Kingery, W. D., Bowen, H. K., and Uhlmann, D. R. (1976). "Introduction to Ceramics," Wiley, New York.]

The random network exhibits local variability of density and structure. The addition of alkali or alkaline earth oxides to the silica glass breaks up the three-dimensional network, and singly bonded oxygen ions are formed that do not participate in the network. The singly bonded oxygen ions are located in the vicinity of the modifying cations to maintain charge balance.

Other glasses have far different structures. For example, glassy B₂O₃ has been shown by X-ray diffraction to be composed of BO₃ triangles. These triangles are linked together in a boroxy configuration (Fig. 4) but distorted with the triangular units linked in ribbons. As alkali or alkaline earth oxides are added to the B₂O₃ boria glass, BO₄ tetrahedra are formed and the structure of the glass takes on a mixed character with BO₃ triangles and BO₄ tetrahedra.

Germania (GeO₂) glass is composed of GeO₄ tetrahedra. Unlike the silica glass, germania glass tetrahedra have a small distribution of bond angles. The essential glassy nature of the material is accommodated by variations in the rotation angle of one tetrahedron to another. Additions of alkali oxide to the germania in small fractions leads to the formation of GeO₆ octahedra in the structure.

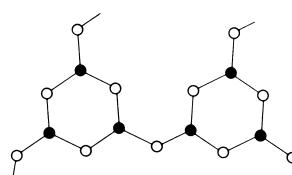


FIGURE 4 Schematic representation of boroxy configurations: ● boron; O, oxygen. [From Kingery, W. D., Bowen, H. K., and Uhlmann, D. R. (1976). "Introduction to Ceramics," Wiley, New York.]

TABLE I Coordination Number and Bond Strength of Oxides^a

	M in MO_x	Valence	Dissociation energy per MO_x (kcal/g atom)	Coordination number	Single-bond strength (kcal/g atom)
Glass formers	B	3	356	3	119
	Si	4	424	4	106
	Ge	4	431	4	108
	Al	3	402–317	4	101–79
	B	3	356	4	89
	P	5	442	4	111–88
	V	5	449	4	112–90
	As	5	349	4	87–70
	Sb	5	339	4	85–68
	Zr	4	485	6	81
Intermediates	Ti	4	435	6	73
	Zn	2	144	2	72
	Pb	2	145	2	73
	Al	3	317–402	6	53–67
	Th	4	516	8	64
	Be	2	250	4	63
	Zr	4	485	8	61
Modifiers	Cd	2	119	2	60
	Sc	3	362	6	60
	La	3	406	7	58
	Y	3	399	8	50
	Sn	4	278	6	46
	Ga	3	267	6	45
	In	3	259	6	43
	Th	4	516	12	43
	Pb	4	232	6	39
	Mg	2	222	6	37
	Li	1	144	4	36
	Pb	2	145	4	36
	Zn	2	144	4	36
	Ba	2	260	8	33
	Ca	2	257	8	32
	Sr	2	256	8	32
	Cd	2	119	4	30
	Na	1	120	6	20
	Cd	2	119	6	20
	K	1	115	9	13
	Rb	1	115	10	12
	Hg	2	68	6	11
	Cs	1	114	12	10

^a From Kingery, W. D., Bowen, H. K., and Uhlmann, D. R. (1976). "Introduction to Ceramics," Wiley, New York.

B. Devitrification

The essential job of the glassmaker is to avoid crystallization of the glass melt as it is cooled to room temper-

ature. If crystals form in the glass due to heat treatment or service at elevated temperatures, the glass is said to have devitrified. In general this is an undesirable event since the crystalline component usually has incompatible

expansion, mechanical, or optical characteristics and leads to degradation of the desired properties such as optical or mechanical. In fact, only certain compositions can be made into glasses without the occurrence of crystalline phases. Compositions of matter that result in satisfactory glasses upon cooling from the melt may form crystalline structures when the equilibrium condition is achieved at some elevated temperature. Materials that form glasses at room temperature have to be cooled relatively quickly into a metastable glass to avoid the crystalline state.

For example, in silica glass, crystals form in the glass if held at even moderately elevated temperatures for long periods of time. The exact conditions that will lead to devitrification are very sensitive to impurity content, especially the alkali oxides. The growth of the crystallites in the glass leads to disastrous consequences such as cracking due to the mismatch of expansion coefficients of the glass and the crystalline form.

However, glass-ceramic is a term given to a class of materials in which partial, controlled devitrification is designed to occur in order to strengthen, harden, and influence the coefficient of thermal expansion of the resultant product. A well-known commercial glass-ceramic is Pyroceram (trademark of Corning Glass Works, Inc.). A glass-ceramic is a two-phase material embodying a crystalline phase in a glass matrix.

Glassy bodies are also made, such as Pyrex (trademark of Corning Glass Works, Inc.), that are composed of

two distinct phases that are both glasses. Such multiphase glasses are also strong and resistant to thermal shock.

IV. COMPOSITION

Natural glasses are formed from magmas that have cooled quickly enough to avoid crystallization and from lightning strikes, the latter known as fulgurites. The fragments of a fulgurite have been found in a sandpit adding up to 9 ft in length and up to 3 in. in diameter. Another group of natural glasses is known as tektites, believed to be of extraterrestrial origin. The compositions of natural glasses vary widely, as illustrated in Table II.

Commercial glass compositions must permit melting at economical temperatures, and a viscosity-temperature relationship that is amenable to forming without devitrification and that results in a glass having the required properties for the specific application intended.

A. Silica

Silica is the glass par excellence having the widest variety of useful applications. It has a low thermal expansion, is resistant to attack by water and acids, and is resistant to devitrification. However, the high temperature required to melt quartz sands, the difficulty in freeing the melt from gas bubbles, and its viscous nature, which makes the

TABLE II Compositions of Some Natural Glasses^a

No.	SiO ₂	Al ₂ O ₃	Fe ₂ O ₃	FeO	MgO	CaO	Na ₂ O	K ₂ O	H ₂ O ⁺	H ₂ O ⁻	TiO ₂
1	97.58	1.54	—	0.23	—	0.38	0.34	—	0.10		
2	87.00	8.00	0.19	1.93	0.82	nil	0.14	0.99	0.36		0.51
3	76.89	12.72	0.43	0.70	0.17	0.57	3.48	4.39	0.47	0.02	0.08
4	76.37	12.59	0.26	0.48	0.17	0.79	3.36	4.67		0.97	0.11
5	74.70	13.72	1.01	0.62	0.14	0.78	3.90	4.02	0.62	—	—
6	73.92	12.38	1.62	0.56	0.27	0.33	3.49	5.39	1.69	—	—
7	71.60	12.44	1.00	0.65	0.06	1.90	3.30	4.22	3.78	0.81	0.25
8	70.62	11.54	1.20	0.18	0.26	1.72	3.52	1.45	7.24	2.42	0.04
9	70.56	20.54	—	0.96	0.11	0.78	3.47	3.38	—	—	—
10	69.95	11.99	0.76	0.64	0.09	0.66	3.70	3.80	4.98	2.80	0.18
11	68.12	12.13	n.d.	1.03	tr.	1.63	5.34	1.69	9.70	—	—
12	67.55	15.68	0.98	1.02	1.11	2.51	4.15	2.86	2.76	0.38	0.34
13	66.68	13.39	0.91	0.21	none	2.72	2.23	2.51	10.05	—	0.38
14	61.4	—	5.8	—	2.9	10.3	2.4	10.1	—	—	—
15	60.12	17.67	3.75	3.40	0.53	2.40	6.78	4.43	0.49	—	—
16	58.59	21.29	4.74	0.71	2.49	6.36	4.42	0.94	0.27	0.04	—
17	53.10	20.70	0.07	4.77	1.77	3.18	9.10	5.84	0.70	—	0.47
18	50.76	14.75	2.89	9.85	6.54	11.05	2.70	0.88	n.d.	—	—

^a From, Morey, G. W. (1954). "The Properties of Glass," Reinhold, New York.

material difficult to form, cause silica glass or fused quartz to be too expensive for most applications. Therefore, other oxides are used to flux (reduce the melting temperature) the quartz raw material. The most effective flux is sodium oxide, Na_2O , but it makes a glass that is soluble in water. Lime, CaO , is added to increase the resistance to moisture (enhance durability) because it is effective in terms of cost and function. Other small constituents occur in the raw ingredients or are added to adjust a host of properties of the glass.

MgO can be added in lieu of CaO . Alumina is added to enhance durability and to lower the coefficient of expansion. Potash (K_2O) substitution for soda (Na_2O) increases durability and reduces the tendency to devitrification. Small amounts of BaO and B_2O_3 are added for the same reason but are relatively expensive.

Not only are the end-use property requirements determinants of the optimum composition, but all the manufacturing demands such as freedom from devitrification and specific forming machine characteristics influence the makeup of the glass batch. The optical glasses have the greatest variety of compositions, because of the many optical glasses needed to fulfill all of the optical designers' needs. **Table III** lists the compositions of a variety of optical glasses.

B. Chemical Durability

The term chemical durability is given to the ability of any specific glass to resist the deteriorating effects of its environments in manufacture or use. These environments include water, carbon dioxide, sulfides, and other atmospheric pollutants, in addition to cleaning and polishing materials used during fabrication of end products. Each corroding substance, as well as various combinations of corrodants, needs to be considered in assessing the suitability of a given composition to a specific application. The nature of the chemical attack is generally not a simple process, because of the complex nature of the glass itself. For example, water corrosion includes the steps of penetration, dissolution of the silicate, and the formation of decomposition products with complex interactions. One cannot describe this process in terms of solubility alone. When the attacking substance includes several ingredients, the chemistry becomes orders of magnitude more complicated. In practical terms, empirical tests are used to define the resistance of a glass to a specific agent.

As an example of such a test for optical glass, polished glass is exposed to 100% relative humidity air and thermally cycled between 46 and 55°C hourly. After having been exposed for various times up to 180 h, the specimens are measured for light scattering. Standard optical

glasses (viz. Schott SK 15, BaF, and F1) undergo surface changes sufficient to lead to light scattering on the order of 2–5% of the incoming beam in a standardized optical test. Schott Glass Technologies, Inc., categorizes climatic resistance into four classes, CR1 to CR4, where CR1 represents glasses with little or no deterioration after 180 h of exposure. The glasses are progressively more susceptible from CR2 to CR4, with CR4 quite prone to deterioration. Under normal humidity conditions, CR1 requires no protection, while the other glasses have to be protected against humid environments.

C. The Surface of Glass

The surface of a glass is different in its nature from the interior of the glass body. This is so because the composition is different as a result of diffusional processes that occur during thermal treatment, of chemical treatments to which the glass has been submitted, and of charge imbalance at the free surface. For example, alkali components diffuse to the surface during heat treatment and evaporate, thus changing the surface concentration of the alkali. Chemical treatments deplete the surface or deposit new material on the surface, either action causing a variation in the surface composition relative to the bulk. Polishing compounds may react with the surface during the polishing process, thus creating chemical gradients. The sorption of gases and vapors from the surrounding environment and the various corrosion processes discussed above serve to alter the surface from the bulk.

V. ANNEALING

A. Viscosity of Glass

When a shearing force is exerted on a liquid, the liquid is displaced in the direction of the force. The rate at which this action occurs is a measure of the viscosity of the liquid. As the viscosity increases, the force required to induce a given rate of displacement increases proportionately. For a unit cube of fluid under shear, the viscosity η is defined by:

$$\eta = Fs/v \quad (\text{dyn/cm}^2)(\text{cm})/\text{cm/sec}$$

where η is viscosity in poise, F is shear force (dynes/cm²), s is the height of the cube (cm), and v is the relative velocity (cm/sec) from the top face to bottom face of cube.

The viscosity of the glass is of importance at all stages of its manufacture. The removal of bubbles and the workability of the glass are critically dependent on viscosity control. **Tables IV** and V present definitions of typical annealing points and strain points of glasses.

TABLE III Compositions of Some Optical Glasses^a

No.	Type	Name			SiO ₂	B ₂ O ₃	Na ₂ O	K ₂ O	CaO	BaO	ZnO	PbO	Al ₂ O ₃	Fe ₂ O ₃
1	496/644	Borosilicate crown	BK3	O802	71.0	14.0	10.0						5.	
2	494/646	Borosilicate crown		O2259	69.59	14	8	3					8	
3	506/602	Crown		O714	74.6		9	11	5					
4	511/640	Borosilicate crown	BKI	O144	70.4	7.4	5.3	14.5	2					
4a	510/640	<i>Same, analysis</i>			72.15	5.88	5.16	13.85	2.04	0	0	Trace	0.04	0.01
5	511/605	Borosilicate crown		O374	68.1	3.5	5.0	16.0				7.0		
6	513/637	Borosilicate crown		O627	68.2	10.0	10.0	9.5				2.0		
7	513/573	Zinc silicate crown	ZK6	O709	70.6		17.0					12.0		
8	516/640	Borosilicate crown	BK7	O3832	69.58	9.91	8.44	8.37	0.07	2.54	0	0	0.04	0.01
9	516/536	Borosilicate crown		O608	53.5	20.0		6.5						
10	516/609	Silicate crown		O40	69.0	2.5	4.0	16.0	8.					
11	517/602	Silicate crown		O60	64.6	2.7	5.0	15.0		10.2	2.0			
12	517/558	Light barium crown		O1092	65.4	2.5	5.0	15.0		9.6	2.0			
13	520/520	High dispersion crown	KF2	O381	66.8		16.0				3.8	11.6	1.5	
13a		<i>Same, analysis</i>			67.40		15.15	0.14	0.39		3.85	10.71	1.72	0.02
14	517/602	Zinc crown	2K2	O546	65.6	4.5	14.5	3.5				11.5		
15	522/596	Silicate crown	K5	O1282	62.5	2.0	5.0	15.0		11.0	3.0	1.0		
16	537/512	Borosilicate flint	KF1	O152	35.4	34.3	7.4					18.7	3.7	
17	540/598	Light barium crown	BaK2	O227	59.5	3.0	3.0	10.0		19.2	5.0			
17a	541/596	<i>Same, analysis</i>			59.13	3.04	3.16	9.70	0.13	19.25	5.0	0	0.11	0.02
18	541/469	Light flint	LLF2	O726	62.6		4.5	8.5				24.1		
19	545/503	Light borosilicate flint		O658	32.75	31	1	3				25	7.	
20	549/461	Extra-light flint	LLF1	O378	59.3		5	8				27.5		
21	553/530	Light barium flint	BaLF5	O846	56.2		1.5	11		15	9	7		
22	563/497	Light barium flint	BaLF1	O543	51.6		1.5	9.5		14	12	11		
23	568/467	Light borosilicate flint		O161	29.0	29.0	1.5					30.0	10.0	
24	568/530	Light barium flint	BaLF3	O602	51.2		5.5	5.0		20	14	4		
25	571/430	Light flint	LF1	O154	54.3	1.5	3.0	8.0				33.0		
25a	571/430	<i>Same, analysis</i>			54.75	0.45	4.31	7.99	0.05	1.64	0.96	29.30	0.04	0.02
26	572/504	Light barium flint		O527	51.7		1.5	9.5		20.0	7.0	10.0		
27	573/574	Light barium crown	BaK1	O211	48.1	4.5	1.0	7.5		28.3	10.1			
27a		<i>Same, analysis</i>			47.73	3.90	1.14	7.16	0.15	29.88	8.61		0.65	0.01
28	576/408	Light flint	LF2	O184	53.7		1.0	8.3				36.6		
29	579/541	Light barium crown		O722	48.8	3.0	0.8	6.5		21.0	15.5	4.1		
29a	580/538	<i>Same, analysis</i>			45.02	4.50	0.64	6.80		22.39	15.53	4.70		0.09
30	581/422	Ordinary light flint	LF3	O276	52.45		4.5	8.0				34.8		
31	583/469	Light barium flint	BaF ₃	O578	49.1		1.0	8.5		13.0	8.5	19.3		
31a	583/463	<i>Same, analysis</i>			49.80		1.24	8.20		13.36	8.03	18.74	0.05	0.01
32	591/605	Dense barium crown	SK5	O2122	37.5	15.0				41.0			5.0	
33	604/438	Barium flint	BaF4	O1266	45.2			7.8		16.0	8.3	22.2		
34	610/574	Densest barium crown	SK1	O1209	34.5	10.1				42.0	7.8		5.0	
34a	610/568	<i>Same, analysis</i>			40.17	5.96	0.13	0.03	0.03	42.35	8.17	None	2.79	0.02
35	612/592	Densest barium crown	SK3	O2071	31.0	12.0				48.0			8.0	
35a	609/588	<i>Same, analysis</i>			34.56	10.96	0.21	0.09		46.91	1.14	None		5.02
36	613/369	Ordinary silicate flint	F3	O118	46.6		1.5	7.8				43.8		

^a From Morey, G. W. (1954). "The Properties of Glass," Reinhold, New York.

TABLE IV Viscosity of Glass^a

	Viscosity (poise)
Flow point	10^5
Softening point	107.6
Upper limit of annealing range	
Incipient softening point	10^{11} – 10^{12}
Upper annealing temperature	
Annealing point (temperature at which glass anneals in 15 min)	$10^{13.4}$
Strain point (glass anneals in ~16 h; below this temperature there is practically no viscous yield)	10^{14}

^a From Musikant, S. (1985). "Optical Materials," Marcel Dekker, New York.

B. Annealing Process

When a glass is cooled down from the melt, strains are introduced due to unequal cooling rates in various parts of the body. These strains can be sufficiently great to introduce severe residual stress in the glass, possibly leading to premature failure. The process of annealing is a thermal treatment designed to reduce these residual strains to an acceptable level for the application intended.

Optical glass requires the most careful annealing because of the significant effect of residual strain on the optical properties of the glass. It is in the development of optical glass that the study of annealing has had its greatest emphasis. Strain causes a homogeneous and isotropic glass to become birefringent. In a birefringent body, the index of refraction is different along the optic axis from the index of refraction in the direction orthogonal to the optic axis. In the case of a plate under a uniaxial load, the optic axis is along the direction of applied load. Remembering that the index of refraction is inversely proportional to the velocity of light in a substance, then the strain is detected by the difference in time of transit of plane-polarized light along the optic axis and transverse to it. This time difference is measured in terms of path difference. The difference is proportional to the strain in the direction of the applied load.

Such measurements are made on an instrument called a polariscope. The stress optical coefficient c is defined as

TABLE V Annealing Point and Strain Point for Typical Glasses^a

	Annealing point (°C)	Strain point (°C)
Borosilicate glass	518–550	470–503
Lime glass	472–523	412–474
Lead glass	419–451	353–380

^a From Musikant, S. (1985). "Optical Materials," Marcel Dekker, New York.

the ratio of the change in optical path difference (OPD) to the unit stress in the colinear direction. As an example, assume:

$$c = 3 \times 10^{-7} \text{ cm/kg cm}^{-2}$$

$$\text{OPD} = 30 \text{ nm or } 30 \times 10^{-7} \text{ cm}$$

$$\text{stress} = 30 \times 10^{-7} / 3 \times 10^{-7}$$

$$= 10 \text{ kg cm}^{-2}$$

Annealing is a two-step process in which the glass is first raised to a sufficiently high temperature that all the strain is removed, and then cooled at a slow enough rate that the permanent strain is low enough to be acceptable. During the annealing process, the glass is heated to a temperature such that the viscosity is sufficiently low that all strains disappear. The glass is then slowly cooled in a controlled manner through the annealing range to the strain point. At that point the glass can be relatively quickly cooled without introducing any additional permanent strain.

A typical allowable birefringence for optical glass is 5 nm/cm path length. A typical annealing curve for an optical glass is shown in Fig. 5.

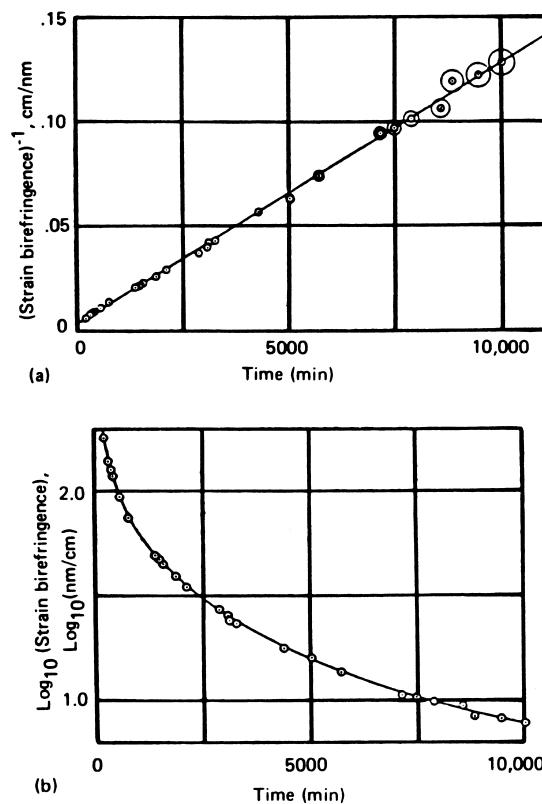


FIGURE 5 Typical annealing curve for glass. (a) Annealing curve, reciprocal of strain birefringence against time, of a glass at 453°C. (b) Annealing curve, logarithm of strain birefringence against time, using the same data as (a). [From Morey, G. W. (1954). "The Properties of Glass," Reinhold, New York.]

TABLE VI Elastic Constants of Glasses at 20°C^{a,b}

Table IX	Fig. 6	Glass	Resonant frequency			Ultrasonic method		
			E	G	ν	E	G	ν
1	A	Silica glass	10.5	4.5	0.165	10.5	4.5	0.17
2	B	96% Silica glass	9.5	4.1	0.16	9.6	4.1	0.18
4	C	Soda-lime plate ^c	10.5	4.3	0.21	10.5	4.35	0.21
6		Soda-lime bulb				10.2	4.15	0.24
8	D	Lead-alkali silicate ^c (50–60% PbO)	7.8	3.2	0.20	8.3	3.35	0.23
10	E	Borosilicate, low-expansion	9.0	3.7	0.22	9.2	3.85	0.20
11		Borosilicate, low-effective-loss				7.3	3.1	0.22
	F	Borosilicate, crown optical ^c	11.6	4.8	0.20	12.0	4.26	0.22

^a From Shand, E. B. (1958). "Glass Engineering Handbook," McGraw-Hill, New York.

^b E, Young's modulus, 10⁶ psi; G, modulus of rigidity, 10⁶ psi; ν, Poisson's ratio.

^c Glasses of same type but not identical for both methods.

VI. PROPERTIES

A. Mechanical Properties

The elastic properties of most interest are the modulus of elasticity (Young's) E, Poisson's ratio ν, and the modulus of rigidity or torsion G. These values are sensitive to the temperature of measurement. For some typical glasses at room temperature, Table VI presents values of these parameters. The effect of temperature on Young's modulus and on the modulus of rigidity are shown in Fig. 6. These numbers were determined by the resonant frequency method, in which a beam in bending is vibrated and the values are computed from the resonant frequency data.

The hardness of various minerals is given in Table VII in terms of Mohs (from scratch tests) and from indentation tests. Table VIII shows the hardness of a number of commercial glasses, the compositions of which are given in Table IX.

The strength of glass varies widely as a function of composition but more importantly as a function of flaw size, flaw distribution, and surface perfection. Glass is a brittle material and fails without yielding. The larger the glass body, the lower the breaking strength, due to the higher probability of the existence of flaws from which fractures can initiate.

Table X presents short-time breaking stress for a variety of annealed glasses tested in flexure. There are many variables, and these values can be considered only as a rough approximation. When designing a structure, the designer must obtain accurate strength data for the specific composition, manufacturing method, shape and size, flaw distribution, surface finish, temperature and environment of use, rate of load application, and any other unique conditions under which the structure must operate.

The extreme variation in strength data for glass fibers is shown in Fig. 7, which shows test data up to 2 million psi for fibers of 0.1 mil (1 mil = 0.001 in.) in diameter

with strengths decreasing to 20,000 psi for fibers of 20 mil diameter.

B. Coefficient of Thermal Expansion

The coefficient of thermal expansion (CTE) of glass is of great importance in the successful application of the

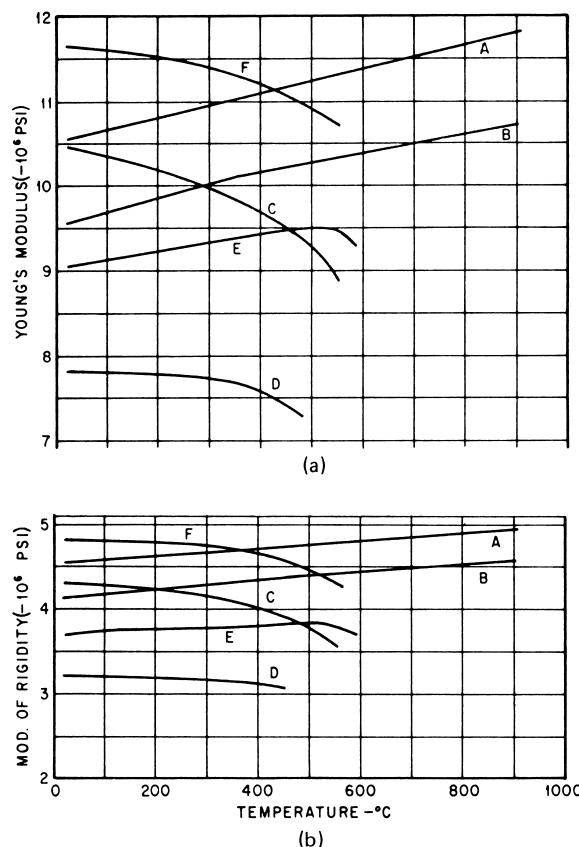


FIGURE 6 Elastic moduli versus temperature. Table VI identifies glasses A–F. [From Shand, E. B. (1958). "Glass Engineering Handbook," McGraw-Hill, New York.]

TABLE VII Hardness of Minerals^a

Mineral	Mohs hardness	Indentation hardness (kg/mm^2)		Bierbaum scratch hardness ^d
		Diamond pyramid ^b	Knoop ^c	
Talc	1	47		1
Gypsum	2	60	32	11
Calcite	3	136	135	129
Fluorite, clear	4	175–200	163–180	143
Apatite	5	659		577
Parallel with optical axis			360	
Perpendicular to optical axis			420	
Albite	6		490	
Orthoclase	6	714	560	975
(Fused silica)		910		
Quartz	7			2700
Parallel with optical axis		1260	710	
Perpendicular to optical axis		1103	790	
Topaz	8	1648	1250	3420
Corundum	9	2085	1655	5300
(Sapphire, synthetic)		2720		
Silicon carbide			2130	
Boron carbide			2265	
Diamond	10		5500–6950	

^a From Shand, E. B. (1958). "Glass Engineering Handbook," McGraw-Hill, New York.^b DPH₅₀, diamond pyramid hardness under 50-g load.^c KHN₅₀, Knoop hardness number under 50-g load.^d Correct., length of impression corrected for resolution limit of microscope.**TABLE VIII Mechanical Hardness of Glasses^a**

Type of glass	Designation in Table IX	Grinding silicon carbide no. 220 grit	Rouge polish	Impact abrasion	Indentation hardness (kg/mm^2)			
					Diamond pyramid		Knoop	
					DPH ₅₀ ^b	Correct. ^d	KHN ₅₀ ^c	Correct. ^d
Silica	1		2.55	3.60	780–800	710–720	640–680	545–575
96% Silica	2			3.53			590	500
Soda-lime window	3		1.00 ^e					
Soda-lime plate	4	1.00 ^e		1.00 ^e –1.07	580	540	575	490
Soda-lime lamp bulb	6	0.98	1.26	1.23	530	500	520	445
Lead-alkali silicate, electrical	7		1.22				430–460	375–400
Lead-alkali silicate, high-lead	8	0.57		0.56	290	270	285–340	250–300
Borosilicate, lowexpansion	10	1.52	1.56	3.10	630	580	550	470
Borosilicate, lowelectrical-loss	11	1.47		4.10				
Aluminosilicate	13	1.36		2.03	640	586	650	550

^a From Shand, E. B. (1958). "Glass Engineering Handbook," McGraw-Hill, New York.^b DPH₅₀, diamond pyramid hardness under 50-g load.^c KHN₅₀, Knoop hardness number under 50-g load.^d Correct., length of impression corrected for resolution limit of microscope.^e Standard for comparison.

TABLE IX Approximate Compositions of Commercial Glasses^{a,b}

No. ^c	Designation	Percent								
		SiO ₂	Na ₂ O	K ₂ O	CaO	MgO	BaO	PbO	B ₂ O ₃	Al ₂ O ₃
1	Silica glass (fused silica)	99.5+								
2	96% Silica glass	96.3	<0.2	<0.2					2.9	0.4
3	Soda-lime—window sheet	71–73	12–15		8–10	1.5–3.5				0.5–1.5
4	Soda-lime—plate glass	71–73	12–14		10–12	1–4				0.5–1.5
5	Soda-lime—containers	70–74		13–16		10–13		0–0.5		1.5–2.5
6	Soda-lime—electric lamp bulbs	73.6	16	0.6	5.2	3.6				1
7	Lead-alkali silicate—electrical	63	7.6	6	0.3	0.2		21	0.2	0.6
8	Lead-alkali silicate—high-lead	35			7.2			58		
9	Aluminoborosilicate (apparatus)	74.7	6.4	0.5	0.9		2.2		9.6	5.6
10	Borosilicate—low-expansion	80.5	3.8	0.4				Li ₂ O	12.9	2.2
11	Borosilicate—low-electrical loss	70.0			0.5			1.2	28.0	1.1
12	Borosilicate—tungsten sealing	67.3	4.6	1.0		0.2			24.6	1.7
13	Aluminosilicate	57	1.0		5.5	12			4	20.5

^a From Shand, E. B. (1958). "Glass Engineering Handbook," McGraw-Hill, New York.

^b In commercial glasses, iron may be present in the form of Fe₂O₃ to the extent of 0.02–0.1% or more. In infrared-absorbing glasses, it is in the form of FeO in amounts from 0.5–1%.

^c Commercial glasses may be identified as follows: 2, Corning Nos. 7900, 7910, 7911, 7912; 6, Corning No. 0080; 7, Corning No. 0010; 8, Corning No. 8870; 9, Kimble N51a; 10, Corning No. 7740; 11, Corning No. 7070; 12, Corning No. 7050; 13, Corning Nos. 1710, 1720.

material. Low-expansion glasses are desirable where even slight dimensional changes are critical in the device, such as in telescope mirror substrates, and where the CTE must match that of an adjacent material such as occurs in glass metal seals of all types.

Composition has a critical influence on the CTE of the glass. Like many other properties of glass, the CTE is roughly an additive factor and for oxide glasses can be approximated by:

$$\lambda = a_1 p_1 + a_2 p_2 + \cdots + a_n p_n$$

TABLE X Breaking Stresses of Annealed Glass, Short-Time Flexure Tests in Air (Effective Duration of Breaking Load, 3 sec)^a

Condition of glass	Avg. breaking stress (psi × 10 ⁻³)
Surfaces ground or sandblasted	1.5–4.0
Pressed articles	3.0–8.0
Blown ware	
Hot iron molds	4.0–9.0
Paste molds	5.0–10.0
Inner surfaces	15.0–40.0
Drawn tubing or rod	6.0–15.0
Polished plate glass	8.0–16.0
Drawn window glass	8.0–20.0
Fine fibers	
Annealed	10.0–40.0
Freshly drawn	30.0–400.0

^a From Shand, E. B. (1958). "Glass Engineering Handbook," McGraw-Hill, New York.

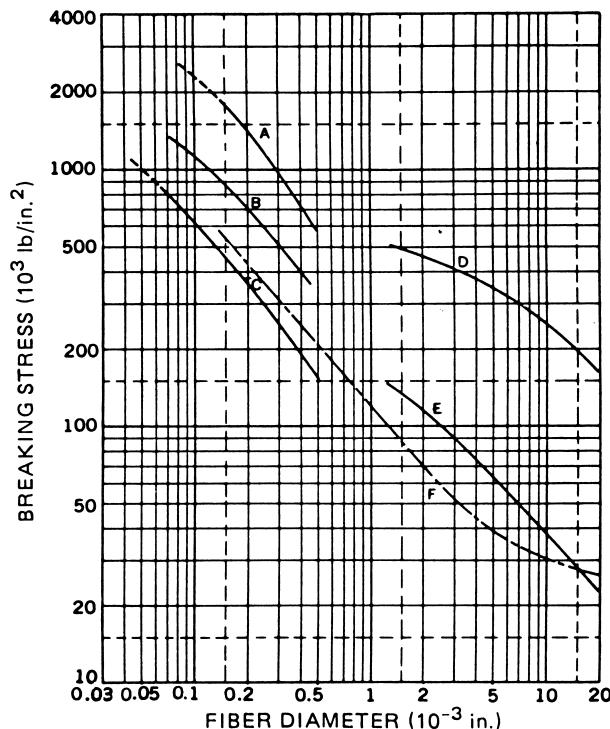


FIGURE 7 Mean tensile strength of glass fibers. A, silica glass in vacuum after baking; B, silica glass in atmosphere dried with CaCl₂; C, silica glass in atmosphere, moistened; D, borosilicate glass in air, acid-fortified; E, borosilicate glass in air; F, soda-lime glass in air. [From Shand, E. B. (1958). "Glass Engineering Handbook," McGraw-Hill, New York.]

TABLE XI Factors for Calculating the Linear Coefficient of Thermal Expansion of Glass^{a,b}

	Winkelmann and Schott	English and Turner	Gillard and Dubrul
SiO ₂	2.67	0.50	0.4
B ₂ O ₃ [†]	0.33	-6.53	-4 + 0.1p
Na ₂ O	33.33	41.6	51 - 0.333p
K ₂ O	28.33	39.0	42 - 0.333p
MgO	0.33	4.5	0
CaO	16.67	16.3	7.5 + 0.35p
ZnO	6.0	7.0	7.75 - 0.25p
BaO	10.0	14.0	9.1 + 0.14p
PbO	10.0	10.6	11.5 - 0.05p
Al ₂ O ₃	16.67	1.4	2

^a From Morey, G. W. (1954). "The Properties of Glass," Reinhold, New York.

^b Values of a_1, a_2, \dots, a_n in $\lambda = a_1 p_1 + a_2 p_2 + \dots + a_n p_n$, in which p_n are the percentages of the given oxides; $\lambda = (10^3/l)(\Delta l/\Delta t)$.

^c Calculated only for that portion of the curve from 0 to 12% B₂O₃.

where λ is the cubical coefficient of expansion, that is, the change in unit volume with change in temperature, p_n is the percentage by weight of the components, and a_n is an empirical factor for each component.

The weighting factors are given in Table XI. The major constituent of most glasses, SiO₂, has a low contribution to the overall CTE, while the alkali metal oxides increase the expansion greatly. Other components are of intermediate nature. However, this type of calculation is most useful in the lower-temperature regime, is very approximate, and must be used with caution.

Figure 8 illustrates linear expansion versus temperature for a number of commercial glasses. Figure 9 illustrates the effect of composition on the CTE for a Na₂O-B₂O₃-SiO₂ glass system. Note the typical knee in the linear expansion curve in Fig. 8. The temperature at which this abrupt change in slope occurs is the glass transition temperature T_g . Below T_g the material is in the glassy state. Between T_g and the melting point, the material is a supercooled liquid.

C. Thermal Conductivity

The thermal conductivity of glass is relatively low. Unlike crystalline materials, in glasses thermal conductivity drops as the temperature decreases. Some representative data are shown in Figs. 10 and 11. For glasses with good infrared transmission, the heat transfer due to radiation becomes significant above about 400°C, in accordance with the radiation heat transfer relations deduced from the Stefan-Boltzmann law:

$$E = kT^4$$

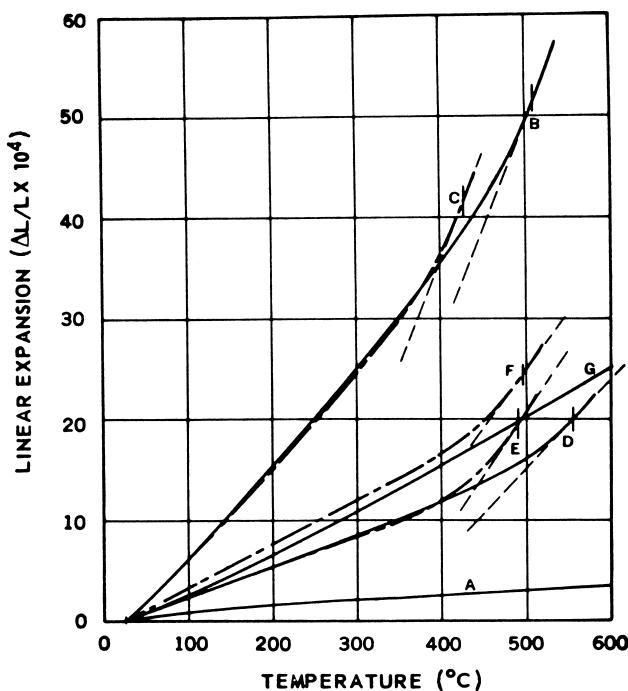


FIGURE 8 Linear expansion of glass with temperature. Light broken lines show increased rates of expansion at annealing points. A, 96% silica glass; B, soda-lime bulb glass; C, medium-lead electrical; D, borosilicate, low expansion; E, borosilicate, low electrical loss; F, borosilicate, tungsten sealing; G, aluminosilicate. [From Shand, E. B. (1958). "Glass Engineering Handbook," McGraw-Hill, New York.]

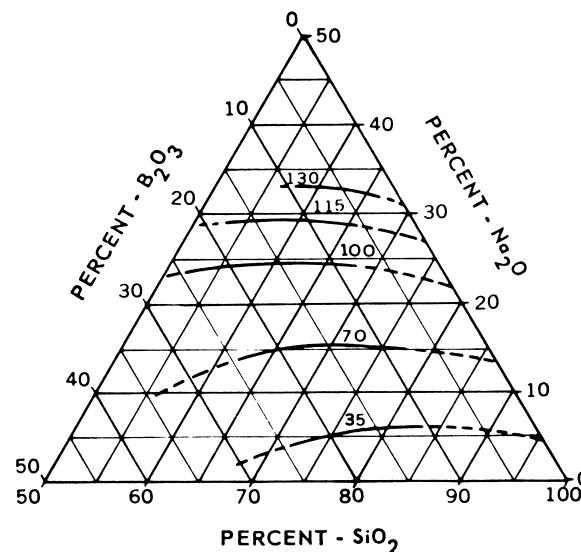


FIGURE 9 Linear coefficients of expansion for the system Na₂O-B₂O₃-SiO₂. Numbers on contours indicate coefficient of expansion per °C × 10⁷. [From Shand, E. B. (1958). "Glass Engineering Handbook," McGraw-Hill, New York.]

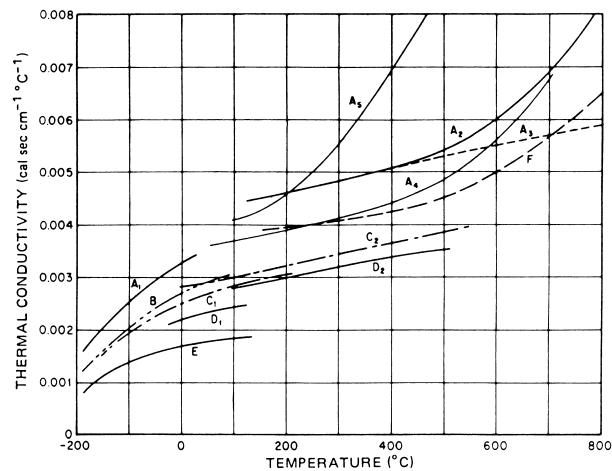


FIGURE 10 Thermal conductivity of glasses. A₁, A₂, A₄, A₅, silica glass; A₃, tin foil on glass surface; B, borosilicate crown glass; C₁, C₂, borosilicate low-expansion glass; D₁, D₂, soda-lime glass; E, lead-alkali silicate glass, 69% PbO; F, 96% silica glass. (Subscripts indicate various sources for data.) [From Shand, E. B. (1958). "Glass Engineering Handbook," McGraw-Hill, New York.]

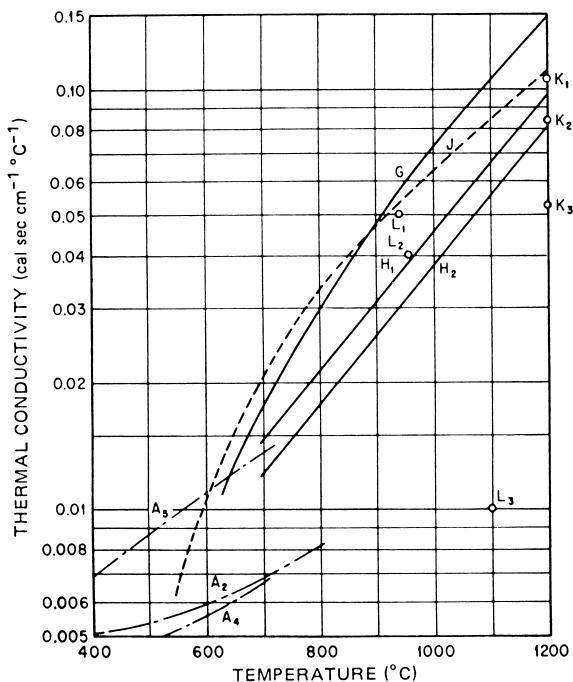


FIGURE 11 Thermal conductivity of glasses in thick sections at high temperatures. G, soda-lime glass; H₁, H₂, soda-lime container glasses; J, low-expansion borosilicate; K₁, soda-lime, 0.015% Fe₂O₃; K₂, soda-lime, 0.16% Fe₂O₃ oxidized; K₃, soda-lime, 0.16% Fe₂O₃ reduced; L₁, soda-lime container glass; L₂, amber glass, 0.2% Fe₂O₃; L₃, pale blue glass, 0.6% Fe₂O₃; A₂, A₄, A₅, see Fig. 10. [From Shand, E. B. (1958). "Glass Engineering Handbook," McGraw-Hill, New York.]

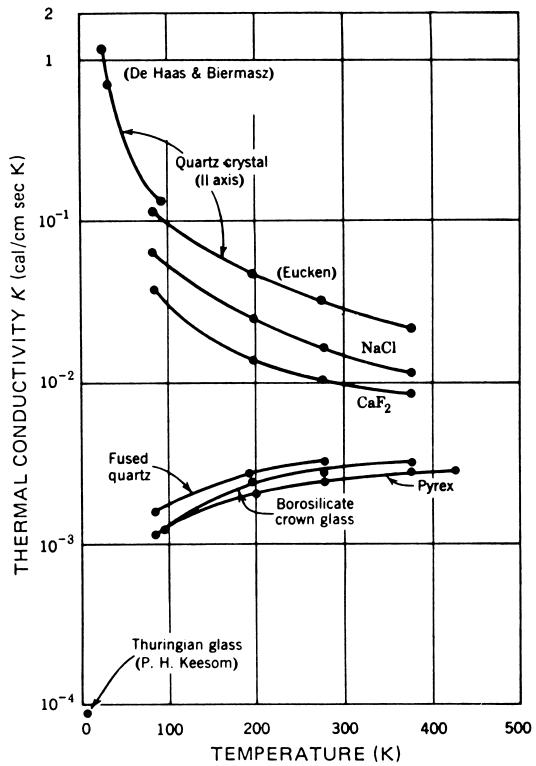


FIGURE 12 Temperature dependence of the thermal conductivity of various crystals and glasses. [From Kittel, C. (1966). "Introduction to Solid State Physics," Wiley, New York.]

where E is the energy per unit time radiating from the surface of black body at absolute temperature T , and K is the proportionality constant.

To convert from the cgs units given in these two charts to English units (Btu/h ft °F), multiply by 241. The conductivity of SiO₂ glass at 50°C is 0.0035 cal/sec cm °C or 0.843 Btu/h ft F.

The thermal conductivity of the quartz crystal at 50°C is approximately 0.025 cal/sec cm °C, in comparison. The temperature dependence of some crystals and glasses is shown in Fig. 12.

D. Specific Heat

Specific heat is defined as the heat energy required to raise the temperature of a unit mass of a substance one unit of temperature. The value c_p indicates that the specific heat is measured at a given pressure p usually 1 atm.

The specific heat of glass can be deduced from measurements of its diffusivity, where

$$\text{thermal diffusivity} = k / \rho c_p$$

where k is thermal conductivity, ρ is density, and c_p = specific heat at constant pressure.

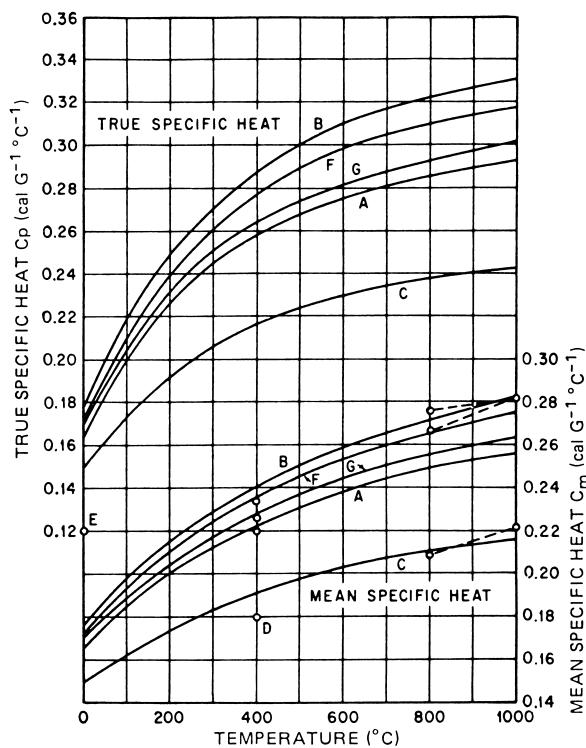


FIGURE 13 Specific heats of glasses. Mean specific heat taken from 0°C. Broken lines from Parmelee and Badger. A, Silica glass (Sosman). B, Soda-lime glass. C, Lead-alkali silicate glass, 22.3% PbO. D, Lead-alkali silicate glass, 26.0% PbO. E, Lead-alkali silicate glass, 46.0% PbO. F, Low-expansion borosilicate glass. G, Aluminosilicate glass. [From Shand, E. B. (1958). "Glass Engineering Handbook," McGraw-Hill, New York.]

Thermal diffusivity is a constant that occurs in the differential equation describing the transient rate of temperature rise in a solid when heat energy is applied to its surface. Diffusivity can be measured by transient thermal tests, and the specific heat c_p can be deduced from such data.

Specific-heat data for a number of glasses are given in Fig. 13.

E. Emissivity

Emissivity of a surface is a measure of its ability to radiate energy in comparison to a black body. For opaque bodies, the relation between reflectivity and emissivity is given by the simple relation

$$E + R = 1$$

where E is emissivity and R is reflectivity.

For bodies transparent to infrared energy, this simple relation becomes more complicated because of radiant energy that emanates from the interior of the body and is transmitted through the body.

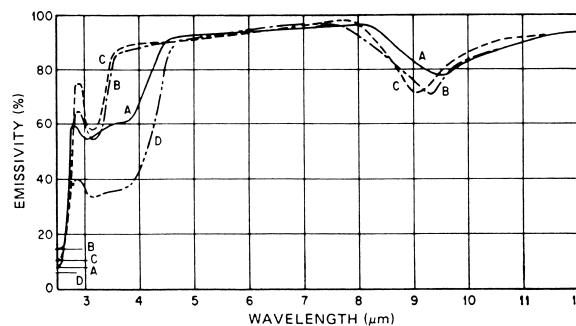


FIGURE 14 Spectral emissivity of glasses. [From Shand, E. B. (1958). "Glass Engineering Handbook," McGraw-Hill, New York.]

Figure 14 illustrates the spectral emissivity of a number of glasses and in addition tabulates the total emissivity for each. Total emissivity is computed by integrating the product of the spectral emissivity and the black body energy over the entire spectrum. As can be seen in the spectral data, the spectral emissivity tends to be high except where the glasses are relatively good transmitters in the near-infrared region.

F. Electrical Properties

Glass is an important material in many electrical devices. Electrical conductivity, dielectric strength, dielectric constant, and dielectric loss are the major electrical properties of interest in most cases.

The electrical conductivity of glasses varies widely, depending on temperature and to a lesser degree on composition. The electrical conductivity largely arises from ionic migration, rather than from free electrons as in the case of metals.

The most significant ions in glasses are the alkali oxide ions, particularly the Na^+ ions. In electrolytic conduction, the ions move from one electrode to the other. If the ions are not replenished, the electrodes become polarized and the conductivity drops off. For dc measurements, sodium amalgams or molten sodium nitrate may be used as electrodes. In ac measurements the polarization problem is minimized.

As the temperature is raised the conductivity rises exponentially and can be expressed as:

$$\sigma = \sigma_0 \exp(-E/RT)$$

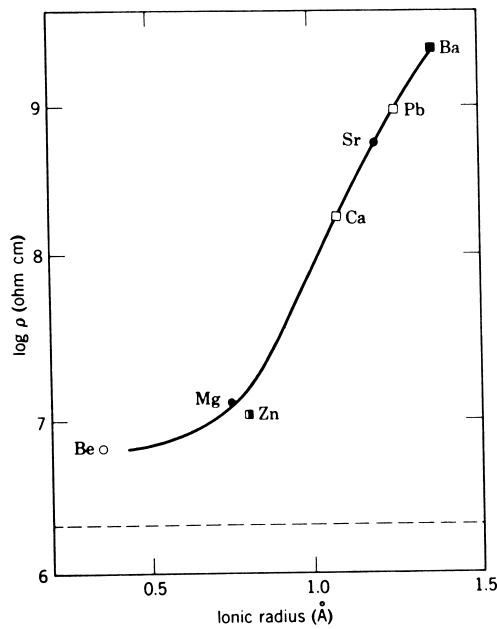


FIGURE 15 Variation of resistivity with divalent ion radius for 0.20 Na₂O–0.20 RO–0.60 SiO₂ glasses. Dashed line is resistivity of 0.20 Na₂O–0.80 SiO₂ glass. [From Mazurin, O. V., and Brailovskii, R. V. (1960). Sov. Phys. Solid State 2, 243.]

where E is an empirical activation energy for conductivity, R is the gas constant, T is the absolute temperature, and σ_0 is a constant.

Compositional changes affect the mobility of the alkali ions and lead to significant changes in conductivity. Figure 15 shows the variation of resistivity with divalent ion radius for 0.20 Na₂O–0.20 RO–0.60 SiO₂ glasses, where R represents a divalent metal atom. The dashed line shows the resistivity of the 0.20 Na₂O–0.80 SiO₂ composition.

Certain oxide glasses do exhibit electronic conductivity, for example, glasses in the vanadium phosphate family. Figure 16 illustrates the exponential increase in conductivity with increasing temperature for a series of V₂O₅–P₂O₅ glasses due to activation of charge carriers.

1. Dielectric Strength

The dielectric strength of a material is a measure of its ability to sustain high-voltage differences without current breakdown. As the voltage across the material is increased, at some value of voltage a burst of current transits through the sample and causes severe damage to the material. The dielectric strength typically is stated in volts per centimeter. Some data for dielectric strength of various glasses are given in Table XII and in Fig. 17. When the dielectric strength is high, the glass will serve as an effective electrical insulator.

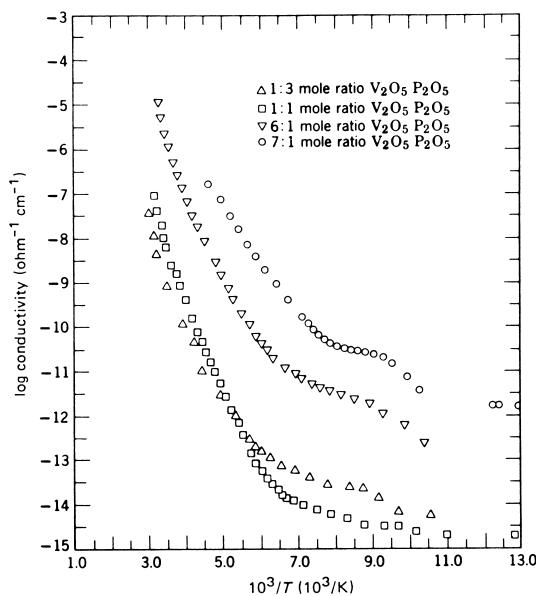


FIGURE 16 Log conductivity as a function of $1/T$ for four typical V₂O–P₂O₅ glasses. [From Schmid, A. B. (1968). J. Appl. Phys. 39, 3140.]

2. Dielectric Constant

The dielectric constant of a material is the ratio of the electrical energy in an electric field within a substance compared to the electrical energy in a similar volume of vacuum subjected to the same electrical field. The value is measured by determining the capacitance of an empty capacitor compared to that of one filled with the material of interest. The dielectric constant is a function of frequency and of temperature.

3. Dielectric Loss Factor

The dielectric loss factor is a measure of the energy absorbed in the medium as an electromagnetic wave passes

TABLE XII Intrinsic Dielectric Strength of Glasses^a

Glass	Dielectric strength (kV/cm, dc)	
	Moon and Norcross	Vermeer
Silica	5000	
Borosilicate, low-expansion	4800	9200
Borosilicate, 16% B ₂ O ₃		11,500
Soda-lime silicate	4500	9000 ^b
Lead-alkali silicate	3100	
Aluminosilicate		9900

^a From Shand, E. B. (1958). "Glass Engineering Handbook," McGraw-Hill, New York.

^b Thuringian glass, 6.2% Al₂O₃.

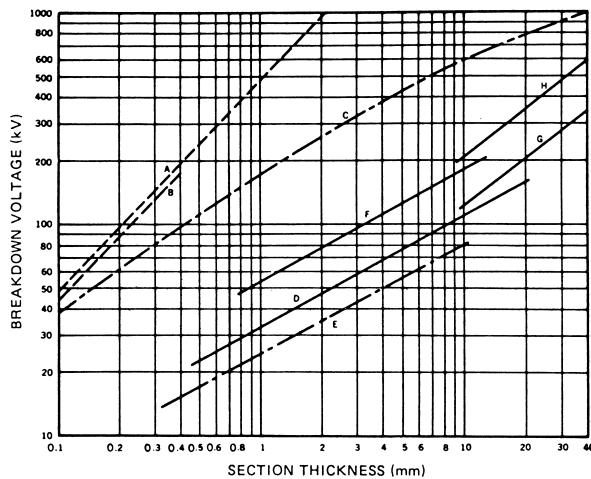


FIGURE 17 Breakdown voltage versus thickness of glass for different conditions at room temperature, 60-cycle voltage raised continuously. A, intrinsic dielectric strength of borosilicate glass; B, intrinsic dielectric strength of soda-lime glass; C, highest test values available for borosilicate glass; D, borosilicate glass plate immersed in insulating oil; E, soda-lime glass plate immersed in insulating oil; F, borosilicate glass plate immersed in semiconducting oil; G, borosilicate glass power-line insulator immersed in insulating oil; H, borosilicate glass power-line insulator immersed in semiconducting oil. [From Shand, E. B. (1958). "Glass Engineering Handbook," McGraw-Hill, New York.]

through that medium. In the ideal case, the losses are zero and the dielectric loss factor is zero. In the case of the capacitor mentioned above, the dielectric loss factor is given by the ratio of the charging current (90° out of phase to the applied voltage) to the loss current in phase with the applied voltage. The total current traversing the condenser is inclined by a power-factor angle between 0 and 90° against the applied voltage v , that is by a loss angle δ , against the $+j$ axis as indicated in Fig. 18. The dissipation factor or $\tan \delta$ is given by I_l/I_c .

Table XIII gives values of dielectric constant and loss tangent for a number of glasses as a function of frequency.

G. Optical Properties

Glass is the great optical material. In its simplest application, the glass has merely to transmit incident visible light such as in a window pane. But for optical devices the glass has to have much more stringent requirements primarily with respect to index of refraction, dispersion (i.e., variation of index of refraction with wavelength), spectral transmittance, absorptance and reflectance, and its scattering characteristics. In addition, the optical glass must have the required mechanical, ther-

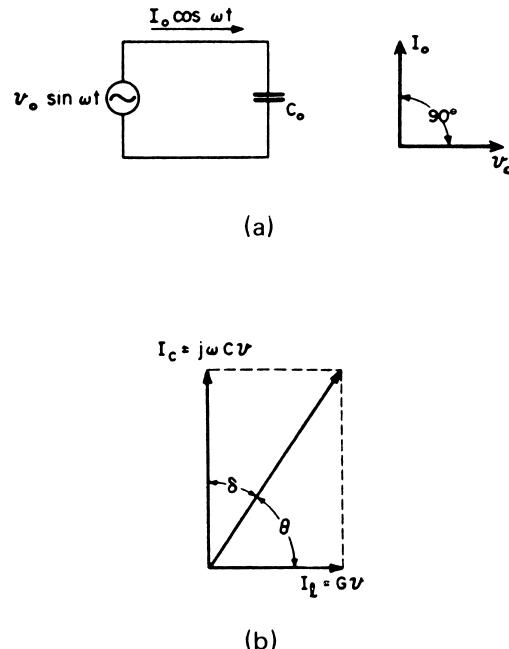


FIGURE 18 (a) Current–voltage relation in ideal capacitor. (b) Capacitor containing dielectric with loss. [From Von Hippel, A. R. (1954). "Dielectric Materials and Applications," M.I.T. Press, Cambridge, MA.]

mophysical, and chemical properties to meet the specific application.

To meet the needs of the wide variety of optical devices, many different optical materials have been developed and are commercially available from optical glass companies around the world. Two of the most significant properties are index of refraction and Abbe number, which is a measure of the dispersion of the glass. The index of refraction is defined as:

$$n = v_0/v$$

where n is the index of refraction, v_0 is the velocity of light in vacuum and v is the velocity of light in the glass.

An index of the dispersion is given by the Abbe number ν_d , defined as:

$$\nu_d = (n_d - 1)/(n_F - n_C)$$

where the various n 's are the indices of refraction of the glass at various spectral lines:

Hydrogen F	4861 Å
Helium d	5876 Å
Sodium D	5893 Å
Hydrogen C	6563 Å

TABLE XIII Properties of Selected Glasses^a (Values for tan δ are Multiplied by 10⁴; Frequency Given in cps)

Glass	t (°C)	Frequency										log ₁₀ Volume resistivity					
		1 × 10 ²	1 × 10 ³	1 × 10 ⁴	1 × 10 ⁵	1 × 10 ⁶	1 × 10 ⁷	1 × 10 ⁸	3 × 10 ⁸	1 × 10 ⁹	2.5 × 10 ¹⁰	2.5 × 10 ¹⁰	25°C	250°C	350°C		
Corning 0010 (potash, soda, lead)	24	ϵ'/ϵ_v	6.68	6.63	6.57	6.50	6.43	6.39	6.33	—	6.1	5.96	5.87	>17	8.9	7.0	
Corning 0014 (lead, barium)	25	ϵ'/ϵ_v	6.78	6.77	6.76	6.75	6.73	6.72	6.70	6.69	—	60	90	110			
Corning 0080 (soda lime)	23	ϵ'/ϵ_v	8.30	7.70	7.35	7.08	6.90	6.82	6.75	—	19.5	—	6.64				
Corning 0090 (potash, lead, silicate)	20	ϵ'/ϵ_v	9.15	9.15	9.15	9.14	9.12	9.10	9.02	—	100	85	90	126	170	180	
Corning 0100 (potash, soda, barium, silicate)	25	ϵ'/ϵ_v	7.18	7.17	7.16	7.14	7.10	7.10	7.07	—	7	18	—	54	103	122	
Corning 0120 (potash, soda, lead)	23	ϵ'/ϵ_v	6.75	6.70	6.66	6.65	6.65	6.65	6.65	—	14	17	24	—	6.64	6.60	
Corning 1770 (soda lime)	25	ϵ'/ϵ_v	6.25	6.16	6.10	6.03	6.00	6.00	6.00	—	12	13	18	—	41	63	
Corning 1990 (iron-sealing glass)	24	ϵ'/ϵ_v	8.40	8.38	8.35	8.32	8.30	8.25	8.20	—	33	26	34	38	—	56	
Corning 1991 (iron-sealing glass)	24	ϵ'/ϵ_v	8.10	8.10	8.08	8.08	8.08	8.06	8.00	—	9	5	7	12	—	84	
Corning 3320 (soda, potash, borosilicate)	24	ϵ'/ϵ_v	5.00	4.93	4.88	4.82	4.79	4.78	4.77	—	43	34	30	32	—	7.99	
Corning 7040 (soda, potash, borosilicate)	25	ϵ'/ϵ_v	4.84	4.82	4.79	4.77	4.73	4.70	4.68	—	34	33	30	32	—	4.74	
Corning 7050 (soda, borosilicate)	25	ϵ'/ϵ_v	4.88	4.84	4.82	4.80	4.78	4.76	4.75	—	50	34	19	22	27	—	
Corning 7052 (soda, potash, lithia, borosilicate)	23	ϵ'/ϵ_v	5.20	5.18	5.14	5.12	5.10	5.10	5.09	—	68	49	24	28	34	—	
Corning 7055	25	ϵ'/ϵ_v	5.45	5.41	5.38	5.33	5.31	5.30	5.27	5.25	—	45	36	30	28	29	38
Corning 7060 (soda, borosilicate)	25	ϵ'/ϵ_v	5.02	4.97	4.92	4.86	4.84	4.84	4.84	—	89	55	40	36	30	—	58
Corning 7070 (potash, lithia, borosilicate)	23	ϵ'/ϵ_v	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	—	50	22	13	10	11	—
Corning 7230 (aluminum borosilicate)	100	ϵ'/ϵ_v	4.17	4.16	4.15	4.14	4.13	4.10	—	—	6	8	11	12	12	12	—
Corning 7570	25	ϵ'/ϵ_v	3.88	3.86	3.85	3.85	3.85	3.85	3.85	—	33	23	16	11	12	—	—
		$\tan \delta$	11.5	13.5	15.9	16.5	19.0	23.5	33	—	11.5	14.54	14.52	14.50	14.42	—	14.2

Continues

TABLE XIII (continued)

Glass	<i>t</i> (°C)	Frequency										\log_{10} Volume resistivity		
		1×10^2	1×10^3	1×10^4	1×10^5	1×10^6	1×10^7	1×10^8	3×10^9	1×10^{10}	2.5×10^{10}	25°C	250°C	350°C
Corning 7720 (soda, lead, borosilicate)	24	ϵ'/ϵ_v	4.74	4.70	4.67	4.64	4.62	4.61	—	—	—	4.59	—	7.2
		$\tan\delta$	78	42	29	22	20	23	—	—	—	43	—	8.8
Corning 7740 (soda, borosilicate)	25	ϵ'/ϵ_v	4.80	4.73	4.70	4.60	4.55	4.52	4.52	—	—	4.52	4.50	8.1
		$\tan\delta$	128	86	65	54	49	45	45	—	—	85	96	6.6
Corning 7750 (soda, borosilicate)	25	ϵ'/ϵ_v	4.45	4.42	4.39	4.38	4.38	—	—	—	—	4.38	4.38	
		$\tan\delta$	45	33	24	20	18	19	—	—	—	43	54	
Corning 7900 (96% silica)	20	ϵ'/ϵ_v	3.85	3.85	3.85	3.85	3.85	3.85	—	3.85	3.84	3.82	3.82	8.1
		$\tan\delta$	6	6	6	6	6	6	—	6	6.8	9.4	13	
	100	ϵ'/ϵ_v	3.85	3.85	3.85	3.85	3.85	3.85	—	3.85	3.84	3.82		
Corning 7911 (96% silica)	25	ϵ'/ϵ_v	—	—	—	—	—	—	—	—	—	—	—	11.7
		$\tan\delta$	—	—	—	—	—	—	—	—	—	—	—	9.6
Corning 8460 (barium, borosilicate)	25	ϵ'/ϵ_v	8.35	8.30	8.30	8.30	8.30	8.30	8.30	—	—	—	—	
		$\tan\delta$	11	9	7.5	7	8	10	10	16	—	40	57	
Corning 8830	25	ϵ'/ϵ_v	5.38	5.28	5.20	5.11	5.05	5.01	5.00	5.00	4.97	—	4.83	
		$\tan\delta$	204	130	91	73	60	54	57	63	—	—	99	
Corning 8871 (alkaline lead silicate)		ϵ'/ϵ_v	8.45	8.45	8.45	8.45	8.45	8.43	—	8.40	8.34	8.05	8.05	7.82
Corning 9010	25	ϵ'/ϵ_v	6.51	6.49	6.48	6.45	6.44	6.43	6.42	6.42	6.40	—	—	6.27
		$\tan\delta$	50.5	36.2	26.7	22.7	21.5	22.6	30	41	—	—	91	
Foanglas (Pittsburgh-Corning) (soda lime)	23	ϵ'/ϵ_v	90.0	82.5	68.0	44.0	17.5	9.0	—	—	—	—	—	5.49
		$\tan\delta$	1500	1600	2380	3200	3180	1960	—	—	—	—	—	4.55
Fused silica 915c	25	ϵ'/ϵ_v	3.78	3.78	3.78	3.78	3.78	3.78	3.78	3.78	3.78	—	—	3.78
		$\tan\delta$	6.6	2.6	1.1	0.4	0.1	0.1	0.3	0.5	—	—	—	1.7
Glass-bonded micas														
Mycalex 400	25	ϵ'/ϵ_v	7.47	7.45	7.42	7.40	7.39	7.38	—	—	—	—	—	7.12
		$\tan\delta$	29	19	16	14	13	13	—	—	—	—	—	33
	80	ϵ'/ϵ_v	7.64	7.59	7.54	7.52	7.50	7.47	—	—	—	—	—	7.32
		$\tan\delta$	150	85	50	25	16	14	—	—	—	—	—	57
Mycalex K10	24	ϵ'/ϵ_v	9.5	9.3	9.2	9.1	9.0	9.0	—	—	—	—	—	11.3 ^b
		$\tan\delta$	170	125	76	42	26	21	—	—	—	40	40	
Mykroy grade 8	25	ϵ'/ϵ_v	6.87	6.81	6.76	6.74	6.73	6.73	6.72	—	—	6.68 ^c	6.96 ^c	6.66
		$\tan\delta$	95	66	43	31	26	24	25	—	—	38	48	81
Mykroy grade 38	25	ϵ'/ϵ_v	7.71	7.69	7.64	7.61	7.61	7.61	—	—	—	7.68 ^c	8.35 ^c	
		$\tan\delta$	43	33	27	24	21	14	—	—	—	35	40	

^a From Gray, D. E. (1963). "American Institute of Physics Handbook," McGraw-Hill, New York. Taken from Tables of Dielectric Materials, Vol. IV, Laboratory for Insulation Research, MIT Technical Report 57; and Properties of Commercial Glasses, Bull. B-83, Corning Glass Works.

^b Not corrected for variations in density.

^c Sample nonhomogeneous.

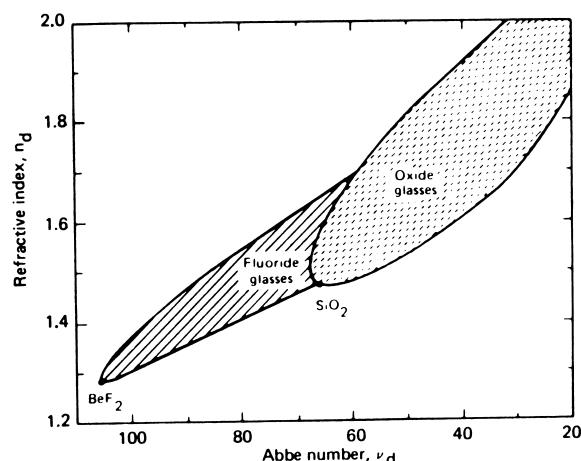


FIGURE 19 For optical materials, n_d (index of refraction) versus v_d (Abbe number). [Adapted from Weber, M. J., Milem, D., and Smith, W. L. (1978). *Opt. Eng.* 17(5), as in Musikant, S. (1985). "Optical Materials," Marcel Dekker, New York.]

The Abbe number is a measure of the chromatic aberration in an optical material. If a single optical element is to be an effective refracting device, then a high index of refraction and a high Abbe number are desirable. The optical glass development of the past century has made available a wide variety of n, v_d combinations. The general range is indicated in Fig. 19.

The other key parameter is transmittance. In general, a high transmittance is required over a specific wavelength or band of wavelengths. Figure 20 illustrates how the infrared transmission is affected by composition of the glass. Glasses composed of higher-atomic-weight elements tend to transmit further into the infrared. This is accompanied by reduced hardness and strength.

Fused quartz glass transmits well in the ultraviolet (UV) portion of the spectrum. However, even very small levels of impurity content severely degrade UV transmittance.

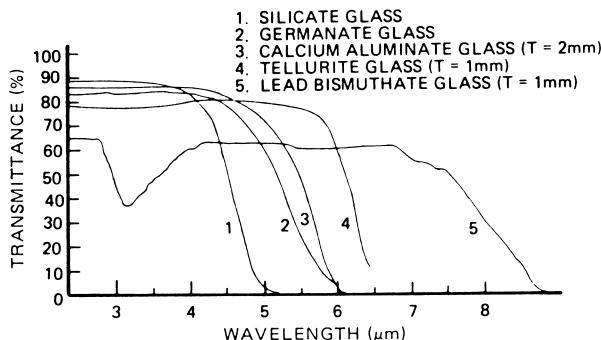


FIGURE 20 Transmittance of infrared-transmitting glasses. [From Dumbaugh, W. H. (1981). *Proc. SPIE* 297.]

During the past decade, the development of optical wave guides in the form of fiber optics has produced glasses of exceedingly high transmittance in wavelengths compatible with the available laser transmitters. Such low-transmittance-loss fibers have made practical long-distance transmission of information over glass-fiber transmission lines.

VII. MANUFACTURE

The manufacture of glass has three essential operations: batching, melting, and forming. In the batching operation, the raw materials are weighed, mixed, and milled as necessary to provide a mixture that can be melted to provide the glass composition desired. Some of the more common raw ingredients are listed in Table XIV. Trace impurities can be a serious problem to the glass maker. One of the most common impurities is iron oxide, which imparts a greenish hue to the glass and prevents heat transfer by radiation through the glass during the melting phase. For high-transmitting fused quartz, extremely high-purity quartz sands are used.

The particle size of the batching materials is a significant variable and must be controlled within close limits to assure a satisfactory melting process. Accuracy of weighing and mixing operations that do not cause segregation in the dry batch are essential to obtaining the homogeneous glass melt desired.

TABLE XIV Glass-Making Materials

Raw material	Chemical composition	Glass-making oxide	Percent of oxide
Sand	SiO ₂	SiO ₂	100.0
Soda ash	Na ₂ CO ₃	NaO ₂	58.5
Limestone	CaCO ₃	CaO	56.0
Dolomite	CaCO ₃ · MgCO ₃	CaO	30.4
		MgO	21.8
Feldspar	K ₂ (Na ₂)O · Al ₂ O ₃ · 6SiO ₂	Al ₂ O ₃	18.0
		K ₂ (Na ₂)O	13.0
		SiO ₂	68.0
Borax	Na ₂ B ₄ O ₇ · 10H ₂ O	Na ₂ O	16.3
		B ₂ O ₃	36.5
Boric acid	B ₂ O ₃ · H ₂ O	B ₂ O ₃	56.3
Litharge	PbO	PbO	100.0
Potash	K ₂ CO ₃ · 1.5H ₂ O	K ₂ O	57.0
Fluorspar	CaF ₂	CaF ₂	100.0
Zinc oxide	ZnO	ZnO	100.0
Barium carbonate	BaCO ₃	BaO	77.7

^a From Shand, E. B. (1958). "Glass Engineering Handbook," McGraw-Hill, New York.

A. Melting

Melting is performed in a gas, oil, or electrically heated furnace. Batch melting in pots or day tanks is used for small quantities of glass. Continuous furnaces ranging from less than 1 ton of glass per day to 1500 tons per day capacity are used in larger production operations. The continuous furnaces are built up from refractory ceramic components and often operate continuously for periods on the order of 1 year before being shut down for rebuild. These furnaces are divided into a large melting section followed by a shallow, narrow refining section (the forehearth) where the glass temperature is reduced preparatory to the forming operation. In fossil-fuel-fired furnaces the hot combustion gases are located above the molten glass. In electric resistance furnaces the glass is directly heated by immersed electrodes. Small induction-heated glass furnaces are used for specialty glasses.

The temperature of the glass in the melting section of the furnace is dependent on the composition. For typical commercial glasses, melting temperatures vary from 1500°C for soda-lime glass to 1600°C for aluminosilicate glass.

B. Refining

The melting and refining processes are very complex. At the cold end of the furnace where the batch is introduced, melting is initiated by the fluxes, which melt first and then dissolve the more refractory ingredients. Elimination of the water vapor and other gases dissolved in the melt takes place in the downstream section of the furnace prior to the entry of the glass into the refining section. In the forehearth the glass is cooled, thereby adjusting its viscosity to the level needed by the forming equipment.

C. Forming

Forming operations include drawing the glass into a sheet, bottle blowing, pressing, tube drawing, rolling into flat glass, and fiber drawing. Fibers can be drawn into continuous strands for textile application or drawn by means of steam or compressed-air jets into fine, discontinuous strands useful for thermal insulation.

After forming, the glass usually has to be annealed to minimize residual strains that could lead to fracture of the glass or, in the case of optical materials, cause unacceptable variations in the optical properties.

D. Finishing

Finally, the ware is subjected to secondary finishing operations such as cutting, grinding, polishing, or thermal or chemical treatments to produce the end item.

VIII. APPLICATIONS

A. Containers

Glass containers must be resistant to the contained fluids, must be sufficiently strong to withstand the internal pressures encountered, as lightweight as possible, and economical to produce. Carbonated beverages can impose pressures up to 125 psi. The containers must be able to withstand the thermal shock imposed by sterilization treatments and impacts that are inflicted during handling and transportation. Careful annealing is performed to minimize residual strain, and quality control is maintained by the use of polariscopic examination.

B. Glazing Glass

Considerations in designing glass for glazing include light transmission, light diffusion (as by sandblasting), absorption of high-energy radiations (as by addition of PbO to absorb X-rays), strength, resistance to thermal gradients and the stresses so induced, heat transfer rate (as by the use of double glazing to inhibit heat transfer), reflectance (either for visible or selective reflectance of infrared), and the ability to heat the glass (as by embedded heating wires for aircraft windshields). These various requirements are met by combinations of composition, heat treatment, coatings, other surface treatments, and by various types of lamination with intervening, transparent films. Bullet-resistant glass is generally made up of multiple layers of glass separated by plastic sheets. The index of refraction of the plastic and the glass have to match closely to assure minimal distortion.

C. Chemical Applications

Laboratory ware and industrial chemical process equipment fabricated from glass are extensively used. Borosilicate glass is widely used because of its relatively low expansion coefficient and reduced susceptibility to thermal shock failure as well as greater chemical durability. However, for special chemical environments other compositions are used, as in boron-free compositions that are resistant to alkaline solutions. Where high-temperature conditions are encountered, fused quartz is used because of its extremely high resistance to thermal shock. Fused quartz is employed in special applications for its excellent ultraviolet transmitting characteristic.

Industrial glass piping is made from heat-resistant glass in a variety of sizes, and various fittings are available. Joints are made by special metal flange arrangements.

D. Lamp Glass

The glass envelope of an electrical lighting device must contain and protect the light-emitting medium, be

TABLE XV Typical Glass Properties^a

Glass type ^b	Description/use ^b	Color ^c	Forms usually available ^e	Mechanical properties			Thermal properties		
				Density (g/cm ³)	Young's modulus ^d 10 ⁶ psi	Poisson's ratio ^e	Impact abrasion resistance ^f	Expansion (10 ⁷ cm/cm ³ /°C) 0-300°C	Thermal endurance ratio ^g
Soda lime 008	General	Clear	T,B,F	2.48	10	0.24	1.2	93	1
Lead 001	Lamp, electronic	Clear	T,B,F	2.81	9	0.21	—	92	1
012	Lamp, electronic	Clear	T,B,F	3.04	8.6	0.22	0.6-0.8	89	1
821	Radiation shielding	Clear	T,B	4.00	—	—	0.6-0.8	88	—
Borosilicate 706	Kovar sealing	Clear	T	2.24	8	0.22	3-4	47	2.2
725	General	Clear	P	2.25	9	0.21	3-4	37	2.5
772	Lead borosilicate	Clear	T,B	2.33	9	0.20	3-4	34	—
776	General	Clear	T,B,P	2.23	9	—	3-4	33	2.5
777	Lead borosilicate	Clear	T	2.30	9	—	3-4	37	—
Aluminosilicate 174	Alkali free, tungsten and molybdenum sealing	Clear	T	2.64	12	0.25	2.0	43	2
177	Alkali free tungsten sealing	Clear	T	2.70	—	—	—	38	—
180	Alkali free, molybdenum sealing	Clear	T	2.74	—	—	—	45	—
Special 250	Encapsulation for capacitors	Clear	Ft.	2.94	—	—	—	60	—
351	Encapsulation for diodes	Light Yellow	Ft.	3.78	—	—	—	44	—
355	Automotive signal	Amber	T,B	2.57	—	—	—	92	—
540	UV transmission	Dark Blue	T,F	2.56	—	—	—	91	—
980	General	Clear	T,B	2.71	10	0.20	—	91	—
982	UV transmission	Clear	T	2.72	10	0.20	—	92	—

Continues

TABLE XV (continued)

Glass type ^b	(°C)	Electrical properties						Optical Properties			
		Viscosity		Electrical resistivity (log ₁₀ ohm cm)		Dielectric constant at 1 MHz		Loss tangent at 1 MHz and 20°C ^j	Loss factor at 1 MHz and 20°C ⁱ	Refractive index ^k <i>n_d</i>	Useful transmittance ^m (nm)
		Strain point (°C)	Anneal point (°C)	250°C	300°C	350°C	and 20°C ^j				
Soda lime											
008	475	515	700	6.2	5.6	5.0	7.2	0.009	0.065	1.512	0.0089
Lead											
001	395	435	625	8.5	7.6	6.6	6.7	0.0015	0.010	1.534	—
012	395	435	625	9.7	8.6	7.6	6.7	0.0014	0.009	1.559	0.0083
821	400	440	595	11.6	10.4	9.4	8.9	0.0005	0.004	1.667	—
Borosilicate											
706	440	485	705	10.0	9.0	8.1	—	—	1.480	—	300-2800
725	505	550	775	7.9	7.1	6.4	4.7	0.003	0.013	1.478	0.0069
772	475	520	760	8.8	7.9	7.2	4.7	0.003	0.013	1.484	0.0076
776	485	535	785	8.5	7.7	7.0	4.5	0.002	0.008	1.471	0.0073
777	485	530	770	9.0	8.1	7.3	—	—	1.480	0.0074	290-3500
Aluminosilicate											
174	690	710	930	12.4	11.5	10.8	—	—	—	—	—
177	805	865	1125	12.2	11.3	10.5	—	—	—	1.522	—
180	745	800	1015	12.7	11.8	11.1	—	—	—	1.536	—
Special											
250	510	545	680	12.3	11.4	10.6	—	—	—	1.568	0.0093
	520	550	638	12.7	11.7	10.8	7.97-8.15	—	—	1.680	0.0323
351	430	475	690	7.5	6.6	5.9	—	—	—	1.511	—
355	440	475	670	7.8	6.8	6.1	—	—	—	—	—
540	470	515	700	9.9	8.9	7.9	—	—	—	1.522	—
980	470	515	695	9.8	8.7	7.8	—	—	—	1.522	—
982	470	515	695	9.8	8.7	7.8	—	—	—	220-4400	—
										220-4400	—

^a From General Electric Co. (1980). "Glass Products," GE Lamp Components Division, Cleveland, OH.^b All forms not always available nor in all sizes.^c T, Tubing and cane; B, blown; Ft., frit or powder; P, pressed; F, formed.^d Data shown in table are estimates to show relative values. They are to be used for reference only.^e Data shown in table are estimates to show relative values. They are to be used for reference only.^f Using soda-lime plate glass as a base of unity, relative values for other glasses are shown.^g Using soda-lime as a base of unity, relative values for other glasses are shown.^h Most glasses are between 0.002 and 0.003 cal/sec (cm °C).ⁱ This is sometimes expressed as a percent (0.009 = 0.9%).^j This is sometimes expressed as a percent (0.065 = 6.5%). Of the standard glasses, lead and borosilicate are the better insulators.^k Values at 589.3 nm.^l Dispersion is shown at $n_f - n_c$.^m Useful transmittance (exceeding 10%) range is shown in nanometers for 1-mm thicknesses. Type 540 has selective transmittance bands. Its useful transmittance is 310-450, 690-1100, and 1100-4600.ⁿ (A) Seldom affected by weathering, (B) could occasionally show weathering effects, and (C) weathering can be a problem.

amenable to hermetic sealing of the electrodes entering and leaving the envelope, be capable of chemically resisting interactions with the substances in the enclosure and external gases at the temperatures of operation, be resistant to thermal shock induced during light-up and at light-off events, be a good electrical insulator, be highly resistant to diffusion of gases, be able to transmit light (or in some cases infrared or ultraviolet energy) with a minimum of transmission losses and with the appropriate degree of scatter, be amenable to mass production methods, and be within the economical cost limits imposed by the market.

To meet the many types of lighting-product requirements, a variety of glasses have been developed by the various manufacturers of glass and lighting devices. **Table XV** lists properties of various glasses, including their useful transmittance ranges, while **Figs. 21, 22**, and **23** show the transmittance spectra for some of these materials.

The most common type of lighting device is the tungsten filament bulb. These bulbs are composed of three parts: the envelope; the stem, and the exhaust tube. For general lighting devices, the envelope is a soda-lime glass, and the stem and exhaust tube are made of a lead-alkali silicate glass. The lead-bearing glasses have a lower softening temperature and a higher electrical resistance than the soda-lime glasses.

The glass bulbs are made by a highly automated glass-blowing process capable of producing several hundred per minute on one machine. Molten glass from the furnace is fed between rollers to produce a ribbon with evenly spaced circular impressions on one face of the glass ribbon. These impressions are next accurately located over an orifice in a steel plate. As the soft glass moves along, the glass sags through the orifice, a blow tip is located at the glass, and the beginning of the envelope caused by the sagging action is completed by a blowing process into a hinged mold, which accurately defines the dimensions of the bulb. After

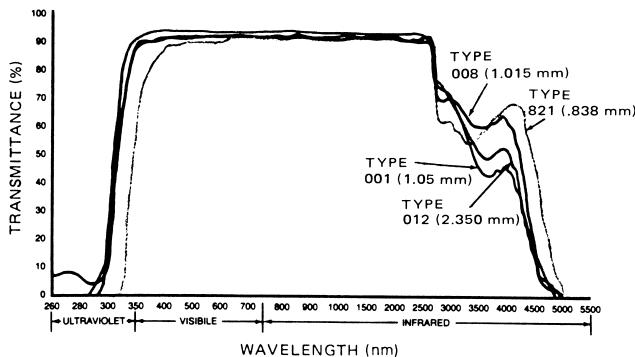


FIGURE 21 Transmittance of soda-lime and lead glasses. Glass thickness shown in parentheses. [From General Electric Co. (1985). "Glass Products," GE Lamp Components Division, Cleveland, OH.]

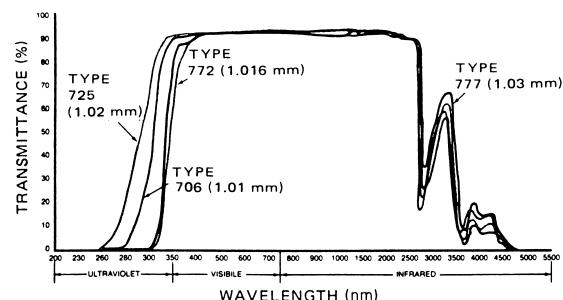


FIGURE 22 Transmittance of borosilicate glasses. [From General Electric Co. (1985). "Glass Products," GE Lamp Components Division, Cleveland, OH.]

cooling, the completed bulb is separated from the ribbon and used in subsequent operations.

Sealing of the metal parts penetrating the glass envelope requires glasses with carefully controlled coefficients of thermal expansion.

E. Glass Fibers

The fact that glass can be drawn into fine fibers has been obvious to every glass maker since ancient times. Therefore, it is surprising that the broad usefulness of glass fibers has not been attained until relatively recent times.

Glass fibers can be thought of in terms of two main categories: (1) continuous strands and (2) staple fibers or short fibers made into mats of various kinds. **Table XVI** displays the various categories of glass fiber properties and applications.

The newest and perhaps most dramatic application for glass fibers is in the field of information transmission via glass-fiber waveguides. The advent of the laser and the associated light detection and electronic signal-processing equipment has made the practical and cost-effective implementation of this technology possible. In addition to information transmission, many new types of instruments are based on the characteristics of coherent light traversing glass-fiber waveguides.

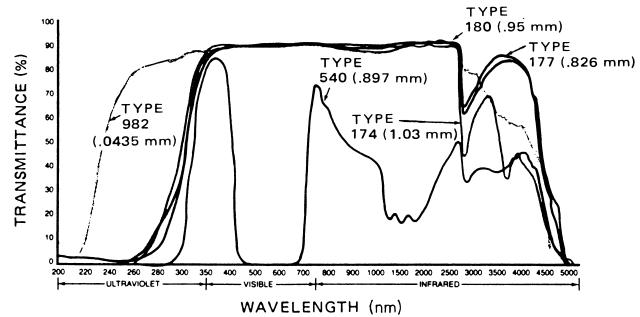


FIGURE 23 Transmittance of aluminosilicate and special glasses. [From General Electric Co. (1985). "Glass Products," GE Lamp Components Division, Cleveland, OH.]

TABLE XVI Properties Related to Applications^a

Glass type	Fibrous-glass forms	Fiber diameter range (in.)	Dominant characteristics	Principal uses
Low-alkali lime-alumina borosilicate	Textiles and mats	0.00023–0.00038	Excellent dielectric and weathering properties	Electrical textiles General textiles Reinforcement for plastics, rubbers, gypsum, papers General-purpose mats
Soda-lime borosilicate	Mats	0.00040–0.00060	Acid resistance	Mats for storage-battery retainers, corrosion protection, water proofing, etc.
	Textiles	0.00023–0.00038		Chemical (acid) filter cloths, anode bags
Soda-lime borosilicate	Wool (coarse)	0.00030–0.00060	Good weathering	Thermal insulations Acoustical products
Soda-lime	Packs (coarse fibers)	0.0045–0.010	Low cost	Coarse fibers only, for air and liquid filters, tower packing, air-washer contact, and eliminator packs
Lime-free soda borosilicate	Wool (fine)	0.00003–0.00020	Excellent weathering	Lightweight thermal insulations, sound absorbers, and shock-cushioning materials
	(Ultrafine)	0.0000 (est.) 0.00003		All-glass high-efficiency filter papers and paper admixtures
High-lead silicate	Textile	0.00023–0.00038	X-ray opacity	Surgical pad strands, X-ray protective aprons, etc.

^a From Shand, E. B. (1958). "Glass Engineering Handbook," McGraw-Hill, New York.

The glass-fiber waveguide must have a low transmittance loss for the wavelength of laser light employed. The waveguide effect is achieved by producing a radial gradient or step in the index of refraction, with the index generally decreasing radially outward. There are many index gradients and step designs used for a variety of purposes. Figure 24 illustrates the application areas and materials used as a function of wavelength. At the present time, silica is the predominant glass fiber used.

IX. THE FUTURE OF GLASS

The statement attributed to Niels Bohr, "Predictions are difficult, especially about the future," holds true in this

instance. However, as long as humans have a progressive technology, that the technology of glass will continue to advance is a simple projection of the past. Certainly each of the application areas mentioned above will continue to develop. As in the case of glass-fiber waveguides, new applications will be discovered in the wake of not yet foreseen new inventions. One thing is certain. The abundance and variety of the raw materials that constitute glass will always provide an economic and technical incentive to use glass as long as the required energy for conversion of the raw stocks is reasonably available.

When humans build their colony on the earth's moon or on the moons of distant planets, the materials for glass making will be there. The asteroid belts will provide the glass-making ingredients for manufacture of structures and other products on the space stations already being designed.

Before recorded history, humans knew glass that had come from outer space. In the not too distant future glass will be made by humans in space.

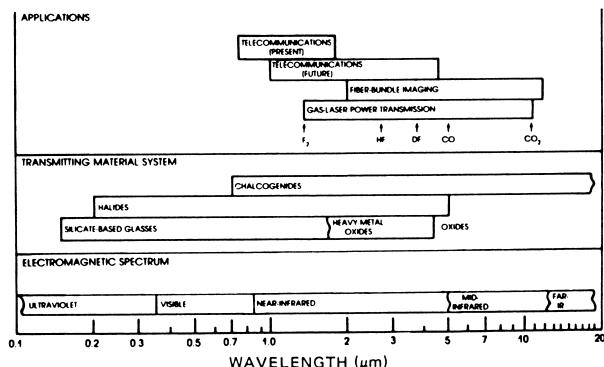


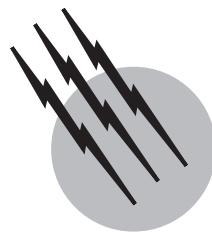
FIGURE 24 Optical fibers: materials and applications. [From Tick, P. A., and Thompson, D. A. (1985). *Photonics Spectra* July.]

SEE ALSO THE FOLLOWING ARTICLES

BONDING AND STRUCTURE IN SOLIDS • CERAMICS • CERAMICS, CHEMICAL PROCESSING OF • GLASS-CERAMICS • LASERS • LIQUIDS, STRUCTURE AND DYNAMICS • OPTICAL FIBER COMMUNICATIONS • PHOTOCHEMOMIC GLASSES

BIBLIOGRAPHY

- Amstock, J. (1997). "Handbook of Glass in Construction," McGraw-Hill, New York.
- ASM International (1991). "ASM Engineered Materials Handbook," Vol. 4: "Ceramics and Glasses," American Society for Metals, Materials Park, OH.
- Barreto, L. M., Kreidl, N., and Vogel, (1994). "Glass Chemistry," 2nd ed., Springer-Verlag, Berlin/New York.
- Bradt, R. C., and Tressler, R. E. (1994). "Fractography of Glass," Kluwar Academic, Dordrecht/Norwell, MA.
- Doremus, R. H. (1994). "Glass Science," 2nd ed., John Wiley & Sons, New York.
- Dumbaugh, W. H. (1981). "Infrared transmitting materials," Proc. SPIE **297**, 80.
- Efimov, A. M. (1996). "Optical Constants of Inorganic Glasses," CRC Press, Boca Raton, FL.
- Gan, F. (1991). "Optical and Spectroscopic Properties of Glass," Springer-Verlag, Berlin/New York.
- General Electric Co. (1980). "Glass Products," GE Lamp Components Division, Cleveland, OH.
- Guillemet, C., and Aben, H. (1993). "Photoelasticity of Glass," Springer-Verlag, Berlin/New York.
- Kingery, W. D., Bowen, H. K., and Uhlmann, D. R. (1976). "Introduction to Ceramics," Wiley, New York.
- Kirsch, R., ed. (1993). "Metals in Glassmaking," Elsevier, Amsterdam/New York.
- Kittel, C. (1966). "Introduction to Solid State Physics," Wiley, New York.
- Kokorina, V. (1996). "Glasses for Infrared Optics," CRC Press, Boca Raton, FL.
- Krause, D., and Bach, H. (1999). "Analysis of the Composition and Structure of Glass and Glass Ceramics," Springer-Verlag, Berlin/New York.
- Morey, G. W. (1954). "The Properties of Glass," Reinhold, New York.
- Musikant, S. (1985). "Optical Materials," Marcel Dekker, New York.
- Neuroth, N., and Bach, H. (1995). "The Properties of Optical Glass," Springer-Verlag, Berlin/New York.
- Sakka, S., Reisfeld, R., and Oehme, (1996). "Optical and Electronic Phenomena in Sol-Gel Glasses and Modern Applications," Springer-Verlag, Berlin/New York.
- Shand, E. B. (1958). "Glass Engineering Handbook," McGraw-Hill, New York.
- Von Hippel, A. R. (1954). "Dielectric Materials and Applications," M.I.T. Press, Cambridge, MA.
- Yamane, M., and Asahara, Y. (2000). "Glasses for Photonics," Cambridge University Press, Cambridge, U.K.
- Zarzycki, J., ed. (1991). "Materials Science and Technology: A Comprehensive Treatment," Vol. 9: "Glasses and Amorphous Materials," John Wiley & Sons, New York.



Glass-Ceramics

Linda R. Pinckney

Corning Incorporated

- I. Design and Processing
- II. Glass-Ceramic Properties
- III. Glass-Ceramic Systems
- IV. Conclusions

GLOSSARY

Ceramming Technical and industrial term used to describe the thermal process by which a glass article is converted to a predominantly crystalline article (i.e., a glass-ceramic).

Metastable phases Phases such as glass or many crystalline structures that are in a state of higher free energy than the thermodynamically stable phase or assemblage of phases.

Phase transformation Transformation from one structural state to another without change in bulk chemical composition (e.g., congruent melting of solids, hexagonal to trigonal quartz).

Primary grain growth Term applied to the growth of crystals in glass from the early stages of nucleation until particle impingement (i.e., the growth of crystals at the expense of the parent glass).

Stuffed derivatives Framework silicates that can be considered as structurally derived from simple polymorphs of silica through substitution of tetrahedral ions of lower valence for silicon (e.g., aluminum, boron) and filling interstitial vacancies with larger cations (e.g.,

lithium, zinc, magnesium, sodium, calcium, potassium) to maintain charge balance.

GLASS-CERAMICS are polycrystalline materials produced by the controlled nucleation and crystallization of glass. Before the discovery in the 1950s that crystallization of glass could be controlled, devitrification, or uncontrolled crystallization, was considered to be a major problem. While devitrification normally results in coarse-oriented crystals containing porosity and results in mechanically weak bodies, glass-ceramics prepared by controlled nucleation are characterized by fine-grained, randomly oriented crystals with some residual glass, no voids or porosity, and generally high strength.

Glass-ceramics can provide significant advantages over conventional glass or ceramic materials by combining the flexibility of forming and inspection with glass with improved and often unique properties in the glass-ceramic. More than \$500 million in glass-ceramic products are sold annually worldwide, ranging from transparent, zero-expansion materials with excellent optical properties and thermal stability to jadelike highly crystalline materials

with excellent strength and toughness. The highest volume is in stovetops and stove windows, architectural cladding, and cookware. Glass-ceramics are also referred to as Pyrocerams (Corning), vitrocerams, and sitalls.

I. DESIGN AND PROCESSING

A. Glass-Ceramic Design

The key variables in the design of a glass-ceramic are glass composition, glass-ceramic phase assemblage, and crystalline microstructure. The glass-ceramic phase assemblage (the types of crystals and the proportion of crystals to glass) is responsible for many of its physical and chemical properties, including thermal and electrical characteristics, chemical durability, and hardness. In many cases these properties are additive; for example, a phase assemblage comprising high- and low-expansion crystals has a bulk thermal expansion proportional to the amounts of each of these crystals.

The nature of the crystalline microstructure (the crystal size and morphology and the textural relationship among the crystals and glass) is the key to many mechanical and optical properties, including transparency-opacity, strength, fracture toughness, and machinability. These microstructures can be quite complex and often are distinct from conventional ceramic microstructures. In many cases, the properties of the parent glass can be tailored for ease of manufacture while simultaneously tailoring those of the glass-ceramic for a particular application.

B. Melting, Forming, and Ceramming

Raw materials such as quartz sand (SiO_2), feldspar [$(\text{Na},\text{K})\text{AlSi}_3\text{O}_8$], and zircon (ZrSiO_4) are commonly employed as batch ingredients. After the batch is mixed thoroughly, it is delivered to the melting furnace, usually in a continuous process, where it goes through several states of progressive refinement before being delivered to the forming process.

Most commercial glass-ceramic products are formed by highly automated glass-forming processes such as rolling, pressing, or casting. The forming method is limited by the liquidus temperature and viscosity of the glass at that temperature. Typical viscosity-temperature curves along with viscosity ranges required for the forming process are shown in Fig. 1. Glasses A and C are compositions based on the β -spodumene solid solution region of the $\text{Li}_2\text{O}-\text{Al}_2\text{O}_3-\text{SiO}_2$ system, and B is representative of $\text{MgO}-\text{Al}_2\text{O}_3-\text{SiO}_2$ glasses, which crystallize predominantly to the cordierite phase. In order to eliminate uncontrolled crystallization, the forming operations must take place at temperatures above the liquidus. Liquidus temperatures for the three glasses in Fig. 1 are 1220°C for A, 1300°C

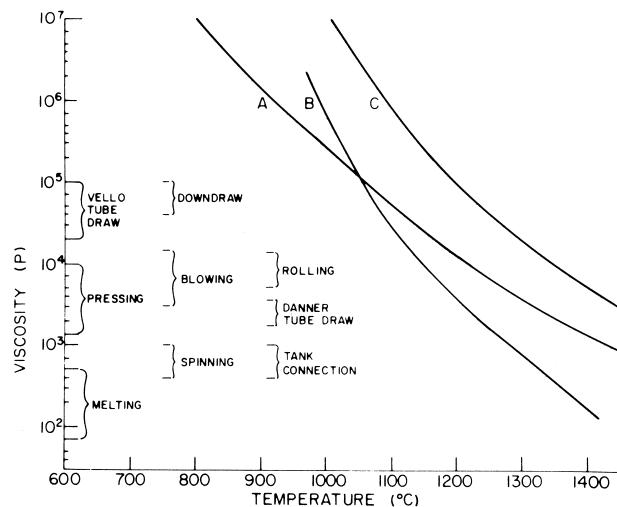


FIGURE 1 Viscosity-forming relationships for three representative glass-ceramic compositions. [After Beall, G. H., and Duke, D. A. (1983). In "Glass: Science and Technology" (D. R. Uhlmann and N. J. Kreidle, eds.), Vol. 1, Academic Press, New York.]

for B, and 1250°C for C. Glass is normally delivered from the tank in a viscosity range of 400 to 1000 P, depending on the desired forming technique. For example, since glass A in Fig. 1 will not crystallize above 1220°C and has a viscosity of 10^4 P at the liquidus temperature, pressing, blowing, and some rolling operations could be performed. Glass C has a viscosity of 5×10^4 P at its liquidus and could therefore be formed by any process except a down-draw. However, casting would be impractical due to the high temperature needed to achieve the required fluidity. Since glass B has a viscosity of only 800 P at its liquidus, forming would be restricted to casting. Thus, it can easily be demonstrated that viscosities and liquidus temperatures of glass-ceramics are significant parameters for using commercial forming processes.

Following the forming process, the glasses are commonly annealed at lower temperatures to remove residual stresses caused by nonuniform cooling. Firepolishing is often used to remove rough or sharp edges resulting from the firepolish, but care must be taken to avoid premature crystallization.

The glass articles are subsequently converted to a crystalline product by the proper heat treatment, known as ceramming. The ceramming process typically consists of a low-temperature hold to induce internal nucleation, followed by one or more higher-temperature holds to promote crystallization and growth of the primary crystalline phase or phases. Although some glass compositions are self-nucleating via homogeneous nucleation, it is more common that certain nucleating agents, such as titanates, metals, or fluorides, are added to the batch to promote phase separation and subsequent internal crystalline nucleation.

Because crystallization occurs at high viscosity, article shapes are usually preserved with little or no (<10%) shrinkage or deformation during the ceramming. In contrast, conventional ceramic bodies typically experience shrinkage of up to 40% during firing. Commercial products manufactured in this manner include telescope mirrors, smooth cooktops, and cookware and tableware.

Glass-ceramics can also be prepared via powder processing methods in which glass frits are sintered and crystallized. Conventional ceramic processing techniques such as spraying, tape- and slip-casting, isostatic pressing, or extrusion can be employed. Such so-called devitrifying frits are employed extensively as sealing frits for bonding glasses, ceramics, and metals. Other applications include cofired multilayered substrates for electronic packaging, matrices for fiber-reinforced composite materials, refractory cements and corrosion-resistant coatings, bone and dental implants and prostheses, architectural panels, and honeycomb structures in heat exchangers.

II. GLASS-CERAMIC PROPERTIES

Given the structural variables of composition, phase assemblage, and microstructure, glass-ceramics can be designed to provide a wide range of physical properties.

A. Thermal Properties

Many commercial glass-ceramics have capitalized on their superior thermal properties, particularly ultralow thermal expansion coupled with high thermal stability and thermal shock resistance—properties that are not readily achievable in glasses or ceramics. Linear thermal expansion coefficients ranging from -7.54 to $20 \times 10^{-6}/^{\circ}\text{C}$ can be obtained. Zero or near-zero expansion materials are used in high-precision optical applications such as telescope mirror blanks as well as for stove cooktops, wood-stove windows, and cookware. Glass-ceramics based on β -eucryptite solid solution display strongly negative thermal expansion (shrinking as temperature increases) and can be valuable for athermalizing precise fiber optic components. At the other extreme, high-expansion devitrifying frits are employed for sealing metals or as corrosion-resistant coatings for metals.

Glass-ceramics have high-temperature resistance intermediate between that of glass and of ceramics; this property depends most on the composition and amount of residual glass in the material. Generally, glass-ceramics can operate for extended periods at temperatures of 700 to over 1200°C . Thermal conductivities of glass-ceramics are similar to those of glass and much lower than those of conventional aluminum oxide-based ceramics. They range from 0.5 to 5.5 W/m K.

B. Optical Properties

Glass-ceramics may be either opaque or transparent. The degree of transparency is a function of crystal size and birefringence, interparticle spacing, and of the difference in refractive index between the crystals and the residual glass. When the crystals are much smaller than the wavelength of light, or when the crystals have low birefringence and the indices of refraction are closely matched, excellent transparency can be achieved.

Certain transparent glass-ceramic materials exhibit potentially useful electrooptic effects. These include glasses with microcrystallites of Cd-sulfoselenides, which show a strong nonlinear response to an electric field, as well as glass-ceramics based on ferroelectric crystals such as niobates, tantalates, or titanates. Such crystals permit electric control of scattering and other optical properties. Transparent glass-ceramics whose crystals can be doped with optically active lanthanide or transition element ions have potential for applications in telecommunications as tunable laser sources or optical amplifiers.

C. Chemical Properties

The chemical durability is a function of the durability of the crystals and the residual glass. Generally, highly siliceous glass-ceramics with low alkali residual glasses, such as glass-ceramics based on β -quartz and β -spodumene, have excellent chemical durability and corrosion resistance similar to that obtained in borosilicate glasses.

D. Mechanical Properties

Like glass and ceramics, glass-ceramics are brittle materials which exhibit elastic behavior up to the strain that yields breakage. Because of the nature of the crystalline microstructure, however, strength, elasticity, toughness (resistance to fracture propagation), and abrasion resistance are higher in glass-ceramics than in glass. Their strength may be augmented by techniques which impart a thin surface compressive stress to the body. These techniques induce a differential surface volume or expansion mismatch by means of ion exchange, differential densification during crystallization, or by employing a lower-expansion surface glaze.

The modulus of elasticity ranges from 80 to 140 GPa in glass-ceramics, compared with about 70 GPa in glass. Abraded modulus of rupture values in glass-ceramics range from 50 to 300 MPa, compared with 40 to 70 MPa in glass. Fracture toughness values range from 1.5 to 5.0 MPa/m in glass-ceramics, compared with less than 1.5 MPa/m in glass. Knoop hardness values of up to 1000 can be obtained in glass-ceramics containing particularly hard crystals such as sapphirine.

E. Electrical Properties

The dielectric properties of glass-ceramics depend strongly on the nature of the crystal phase and on the amount and composition of the residual glass. In general, glass-ceramics have such high resistivities that they are used as insulators. Even in relatively high alkali glass-ceramics, alkali migration is generally limited, particularly at low temperatures, because the ions are either incorporated into the crystal phase or they reside in isolated pockets of residual glass. Nevertheless, by suitably tailoring the crystal phase and microstructure, reasonably high ionic conductivities can be achieved in certain glass-ceramics, particularly those containing lithium.

Glass-ceramic loss factors are low, generally much less than 0.01 at 1 MHz and 20°C. Glass-ceramics comprised of cordierite or mica crystals, for example, typically provide loss factors less than 0.001. The fine-grained, homogeneous, nonporous nature of glass-ceramics also gives them high dielectric breakthrough strengths, especially compared with ceramics, allowing them to be used as high-voltage insulators or condensors.

Most glass-ceramics have low dielectric constants, typically 5–8 at 1 MHz and 20°C. Glass-ceramics comprised primarily of network formers can have dielectric constants as low as 4, with even lower values ($K < 3$) possible in microporous glass-ceramics. On the other hand, very high dielectric constants of over 1000 can be obtained from relatively depolymerized glasses with crystals of high dielectric constant, such as lead or alkaline earth titanate.

III. GLASS-CERAMIC SYSTEMS

All commercial as well as most experimental glass-ceramics are based on silicate bulk glass compositions, although there are numerous nonsilicate and even nonoxide exceptions. Glass-ceramics can be further classified by the composition of their primary crystalline phases, which may consist of silicates, oxides, phosphates, borates, or fluorides. Examples of commercial glass-ceramics are given in [Table I](#).

A. Simple Silicates

The most important simple silicate glass-ceramics are based on lithium metasilicate Li_2SiO_3 , lithium disilicate $\text{Li}_2\text{Si}_2\text{O}_5$, diopside $\text{CaMgSi}_2\text{O}_6$, wollastonite CaSiO_3 , and enstatite MgSiO_3 .

There are four groups of commercially important lithium silicate glass-ceramics, three of which are nucleated with P_2O_5 . The first comprises strong, fine-grained materials based on a microstructure of fine-grained lithium

disilicate crystals with dispersed nodules of quartz or cristobalite crystals. Several companies have introduced these finely textured glass-ceramics for use as magnetic disk substrates for portable computer hard drives. More recently, internally nucleated glass-ceramics in the $\text{SiO}_2\text{-Li}_2\text{O}\text{-K}_2\text{O}\text{-ZnO}\text{-P}_2\text{O}_5$ system based on lithium disilicate with high mechanical strength and excellent chemical durability have been developed for use as dental overlays, crowns, and bridges. The third group comprises high-expansion lithium disilicate glass-ceramics which match the thermal expansion of several nickel-based superalloys and are used in a variety of high-strength hermetic seals, connectors, and feedthroughs.

The fourth group is nucleated with colloidal silver, gold, or copper, which in turn is photosensitively nucleated. By suitably masking the glass and then irradiating with ultraviolet light, it is possible to nucleate and crystallize only selected areas. The crystallized portion consists of dendritic lithium metasilicate crystals, which are much more soluble in dilute hydrofluoric acid than is the glass. The crystals can thus be etched away, leaving the uncrystallized (masked) portion intact. The resulting photoetched glass can then be flood-exposed to ultraviolet rays and heat-treated at higher temperature, producing the stable lithium disilicate and quartz phases. The resulting glass-ceramic is strong, tough, and faithfully replicates the original photoetched pattern. These chemically machined materials have been used as fluidic devices, lens arrays, magnetic recording head pads, and charged plates for ink-jet printing.

Blast furnace slags, with added sand and clay, have been used in Eastern Europe for over 30 years to manufacture inexpensive nonalkaline glass-ceramics called slagsitall. The primary crystalline phases are wollastonite and diopside in a matrix of aluminosilicate glass. Metal sulfide particles serve as nucleating agents. The chief attributes of these materials are high hardness, good to excellent wear and corrosion resistance, and low cost. The relatively high residual glass levels (typically $>30\%$), coupled with comparatively equiaxial crystals, confers only moderately high mechanical strengths of ~ 100 MPa. Slagsitall materials have found wide use in the construction, chemical, and petrochemical industries.

More recently, attractive translucent architectural panels of wollastonite glass-ceramics have been manufactured by Nippon Electric Glass and sold under the trade name Neopariés. A sintered glass-ceramics with about 40% crystallinity, this material can be manufactured in flat or bent shapes by molding during heat treatment. It has a texture similar to that of marble but with greater strength and durability than granite or marble. Neopariés is used as a construction material for flooring and for exterior and interior cladding.

TABLE I Commercial Glass-Ceramic Compositions (wt%)^a

	(A)	(B)	(C)	(D)	(E)	(F)	(G)
SiO ₂	68.8	55.5	63.4	69.7	56.1	47.2	60.9
Al ₂ O ₃	19.2	25.3	22.7	17.8	19.8	16.	14.2
Li ₂ O	2.7	3.7	3.3	2.8	—	—	—
MgO	1.8	1.0	^b	2.6	14.7	14.5	5.7
ZnO	1.0	1.4	1.3	1.0	—	—	—
BaO	0.8	—	2.2	—	—	—	—
P ₂ O ₅	—	7.9	^b	—	—	—	—
CaO	—	—	—	—	0.1	—	9.0
Na ₂ O	0.2	0.5	0.7	0.4	—	—	3.2
K ₂ O	0.1	—	^b	0.2	—	9.5	1.9
Fe ₂ O ₃	0.1	0.03	^b	0.1	0.1	—	2.5
MnO	—	—	—	—	—	—	2.0
B ₂ O ₃	—	—	—	—	—	8.5	—
F	—	—	—	—	—	6.3	—
S	—	—	—	—	—	—	0.6
TiO ₂	2.7	2.3	2.7	4.7	8.9	—	—
ZrO ₂	1.8	1.9	1.5	0.1	—	—	—
As ₂ O ₃	0.8	0.5	^b	0.6	0.3	—	—
Primary phases	β-Quartz	β-Quartz	β-Quartz	β-Spodumene	Cordierite	Fluormica cristobalite	Diopside

^a Commercial applications and sources:

- (A) Transparent cookware (Visions); World Kitchen.
- (B) Telescope mirrors (Zerodur); Schott Glaswerke.
- (C) Infrared transmission cooktop (Ceran); Nippon Electric Glass.
- (D) Cookware, hot-plate tops; World Kitchen.
- (E) Radomes; Corning Incorporated.
- (F) Machinable glass-ceramic (Macor); Corning Incorporated.
- (G) Gray slagsitall; Hungary.

^b Not available.

B. Fluorosilicates

Compared to the simple silicates, fluorosilicate crystals have more complex chain and sheet structures. Examples from nature include hydrous micas and amphiboles, including hornblende and nephrite jade. In glass-ceramics, fluorine replaces the hydroxyl ion; fluorine is much easier to incorporate into glass and also makes the crystals more refractory.

1. Sheet Fluorosilicate Glass-Ceramics

Glass-ceramics based on sheet silicates of the fluorine mica family can be made machinable and strong, with excellent dielectric properties and thermal shock resistance. This combination of properties stems from their unique microstructure of interlocking and randomly oriented flakes of cleavable mica crystals. Because micas can be easily delaminated along their cleavage planes, fractures propagate readily along these planes but not along

other crystallographic planes. As a result, the random intersections of the crystals in the glass-ceramic cause crack branching, deflection, and blunting, thereby arresting crack growth (Fig. 2). In addition to providing the material with high intrinsic mechanical strength, the combination of ease of fracture initiation with almost immediate fracture arrest enables these glass-ceramics to be readily machined. Flexural strengths, in the range of 135–175 MPa, are inversely proportional to the flake diameter and are not sensitive to abrasion.

The commercial glass-ceramic Macor (Corning Code 9658), based on a fluorophlogopite mica, $KMg_3AlSi_3O_{10}F_2$, is capable of being machined to high tolerance (± 0.01 mm) by conventional high-speed metalworking tools. By suitably tailoring its composition and nucleation temperature, relatively large mica crystals with high two-dimensional aspect ratios are produced, enhancing the inherent machinability of the material. The growth of high-aspect-ratio flakes can be enhanced by designing the composition to (1) delay nucleation of



FIGURE 2 Crack deflection by mica crystals in fluorophlogopite glass-ceramics. [After Grossman, D. G. (1977), *Vacuum* **28**(2), 55–61.]

phlogopite until relatively high temperatures where growth is rapid, (2) stimulate lateral platy growth as opposed to thickening by limiting the concentration of the cross-bonding species potassium, and (3) produce a relatively fluid B_2O_3 -rich residual glass allowing rapid diffusion of species to the edges of the growing crystals. This “house-of-cards” microstructure is illustrated in Fig. 3.

In addition to precision machinability, Macor glass-ceramic has high dielectric strength, very low helium permeation rates, and is unaffected by radiation. This glass-ceramic has been employed in a wide variety of applications including high-vacuum components and hermetic joints, precision dielectric insulators and components, seismograph bobbins, sample holders for field ion microscopes, boundary retainers for the space shuttle, and gamma-ray telescope frames.



FIGURE 3 Microstructure of crystallized fluorophlogopite glass-ceramics (white bar = 1 μm). [After Grossman, D. G. (1977), *Vacuum* **28**(2), 55–61.]

Translucent machinable glass-ceramics also have been developed for use in restorative dentistry. They are based on the tetrasilicic fluormica $KMg_{2.5}Si_4O_{10}F_2$. High strength (~140 MPa) and low thermal conductivity along with a hardness closely matching natural tooth enamel provide this material with advantages over conventional metal ceramic systems.

Since the mid-1980s, the inherent strength and machinability of mica-based materials have extended the growing field of biomaterials. Biomaterials for bone implants demonstrate biocompatibility—are well tolerated by the body—and may even offer bioactivity, the ability of the biomaterial to bond with the hard tissue (bones) of the body. Several strong, machinable, and biocompatible mica- and mica/fluorapatite glass-ceramics have been widely studied as bone implants. One, known as Bioverit I, is based on a microstructure of tetrasilicic fluormica, $KMg_{2.5}Si_4O_{10}F_2$, and fluorapatite, $Ca_{10}(PO_4)_6F_2$. The mica phase imparts strength and machinability to the material while the fluorapatite confers bioactivity. The glass-ceramics have long-term stability and are machinable with standard metal-working tools. Another glass-ceramic, Bioverit II, possesses an unusual microstructure of curved fluormica crystals with a “cabbage-head” morphology, with cordierite ($Mg_2Al_4Si_5O_{18}$) as an accessory phase. The analyzed composition of the curved mica crystals shows them to be slightly enriched in alumina: $(Na_{0.18}K_{0.82})(Mg_{2.24}Al_{0.61})(Si_{2.78}Al_{1.22})O_{10.10}F_{1.90}$. While this material is not bioactive, it is strong, highly machinable, and very well tolerated biologically.

2. Chain Fluorosilicate Glass-Ceramics

Interlocking blade- or rodlike crystals can serve as important strengthening or toughening agents, much as fiber glass is used to reinforce polymer matrices. Naturally occurring, massive aggregates of chain silicate amphibole crystals, such as nephrite jade, are well known for their durability and high resistance to impact and abrasion. Glass-ceramics with microstructures of randomly oriented, highly anisotropic chain silicate crystals generally provide superior strength and toughness, for in order for a fracture to propagate through the material, it generally will be deflected and blunted as it follows a tortuous path around or through cleavage planes of each crystal. Indeed, glass-ceramics based on chain silicate crystals have the highest toughness and body strength of any glass-ceramics.

Glass-ceramics based on the amphibole potassium fluorrichterite ($KNaCaMg_5Si_8O_{22}F_2$) have a microstructure consisting of tightly interlocked, fine-grained, rod-shaped amphibole crystals in a matrix of minor cristobalite, mica, and residual glass (Fig. 4). The flexural strength of these

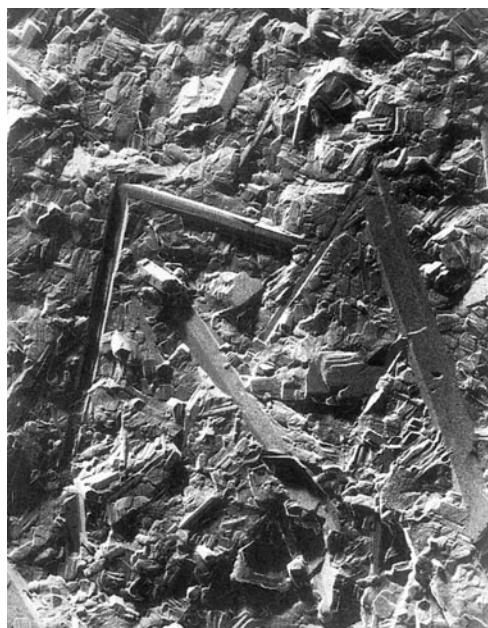


FIGURE 4 Fracture surface of fluorricherite glass-ceramic showing the effects of rod reinforcement toughening.

materials can be further enhanced by employing a compressive glaze. Richterite glass-ceramics have good chemical durability, are usable in microwave ovens, and, when glazed, resemble bone china in their gloss and translucency. These glass-ceramics have been manufactured for use as high-performance institutional tableware and as mugs and cups for the Corelle (Corning) line.

An even stronger and tougher microstructure of interpenetrating acicular crystals is obtained in glass-ceramics based on fluorcanasite ($K_{2-3}Na_{4-3}Ca_5Si_{12}O_{30}F_4$) crystals. Cleavage splintering and high-thermal-expansion anisotropy augment the intrinsic high fracture toughness of the chain silicate microstructure in this highly crystalline material. Canasite glass-ceramics are being evaluated as potential dental materials.

C. Aluminosilicate Glass-Ceramics

Aluminosilicates consist of frameworks of silica and alumina tetrahedra linked at all corners to form three-dimensional networks; familiar examples are the rock-forming minerals quartz and feldspar. Commercial glass-ceramics based on framework structures comprise compositions from the $Li_2O-Al_2O_3-SiO_2$ (LAS) and the $MgO-Al_2O_3-SiO_2$ (MAS) systems. The most important, and among the most widespread commercially, are glass-ceramics based on the various structural derivatives of high (β)-quartz and on β -spodumene and cordierite. These silicates are important because they possess very

low bulk thermal expansion characteristics, with the consequent benefits of exceptional thermal stability and thermal shock resistance. Thus materials based on these crystals can suffer large thermal upshock or downshock without experiencing strain that can lead to rupture, a critical property in products such as stovetops, missile nose cones, and cookware. Representative aluminosilicate glass-ceramic compositions are given in **Table I**.

Although glass-ceramics with low to moderate thermal expansion can be formed from a wide variety of glasses in the $Li_2O-Al_2O_3-SiO_2$ system, the most useful are based on solid solutions of β -quartz and β -spodumene. These are highly crystalline materials with only minor residual glass or accessory phases. Partial substitution of MgO and ZnO for Li_2O improves the working characteristics of the glass while lowering the materials cost. Glass-ceramics containing β -quartz or β -spodumene can be made from glass of the same composition by modifying its heat treatment: β -quartz is formed by ceramming at or below $900^\circ C$, and β -spodumene by ceramming above $1000^\circ C$.

1. β -Quartz Solid-Solution Glass-Ceramics

Transparent glass-ceramics with near-zero thermal expansion are obtained through the precipitation of β -quartz solid solution. A mixture of ZrO_2 and TiO_2 produces highly effective nucleation of β -quartz, resulting in very small (<100-nm) crystals. This fine crystal size, coupled with low birefringence in the β -quartz phase and closely matched refractive indices in the crystals and residual glass, results in a transparent yet highly crystalline body.

Glass-ceramics with ultralow thermal expansion are particularly valuable in applications for which thermal dimensional stability is critically important. The best-known low-expansion optical material is the glass-ceramic Zerodur, manufactured by Schott. Its composition was designed specifically for maximum temperature-time stability of the metastable β -quartz phase and for constant thermal expansion characteristics within this range. Ambient temperature changes from $-50^\circ C$ to $+50^\circ C$ produce length changes of only a few parts per million. This property is critical for applications such as telescope mirrors, in which large volumes are cast, to ensure constant thermal expansion over the entire volume. This glass-ceramic can be readily polished to optical quality as required for reflective mirrors. Zerodur is also used in reflective optics and as the base block for ring laser gyroscopes, which have generally replaced mechanical gyros in airplanes.

Another important commercial application for β -quartz glass-ceramics with a low coefficient of thermal expansion is the smooth, radiant cooktop for electric stoves, as seen in **Fig. 5**. In this case, the glass-ceramic typically is doped with 0.1% V_2O_5 in order to render the material black in



FIGURE 5 Radiant stovetop composed of nanocrystalline β -quartz solid-solution glass-ceramic with near-zero coefficient of thermal expansion. [Photograph courtesy of Eurokera, S.N.C.]

general appearance. The vanadium does, however, allow transmission in the red and near-infrared, where the tungsten halogen lamp and other resistive elements radiate energy. **Figure 6** illustrates the transmission of a Corning cooktop versus the radiation from a typical tungsten halogen lamp. While the glass-ceramic cuts out most visible radiation, it transmits very effectively in the red and near-infrared. Other consumer applications for transparent stuffed β -quartz glass-ceramics include transparent cookware, woodstove windows, and fire doors.

2. β -Spodumene Solid-Solution Glass-Ceramics

Opaque, low-expansion glass-ceramics are obtained by ceramming these LAS materials at temperatures above 1000°C. The transformation from β -quartz to β -spodumene takes place between 900 and 1000°C and is accompanied by a 5- to 10-fold increase in grain size (to 1–2 μm). When TiO_2 is used as the nucleating agent, rutile development accompanies the silicate phase transformation. The combination of larger grain size with

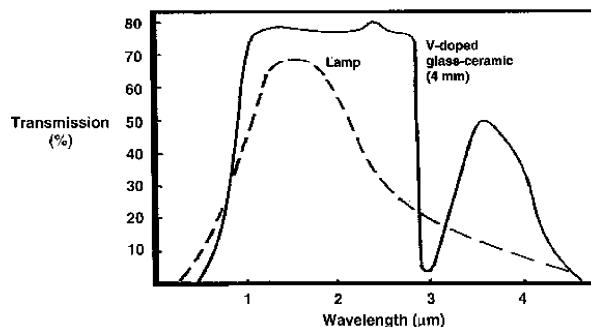


FIGURE 6 Transmission of a Corning glass-ceramic cooktop versus tungsten-halogen lamp.

the high refractive index and birefringence of rutile gives the glass-ceramic a high degree of opacity. Secondary grain growth is sluggish, giving these materials excellent high-temperature dimensional stability. The stable solid solution of β -spodumene varies in composition of $\text{Li}_2\text{O}\text{-}\text{Al}_2\text{O}_3\text{-SiO}_2$ from 1:1:4 to about 1:1:10. **Figures 7** and 8 show the low thermal expansion characteristics of these tetragonal crystals in the range from 0 to 1000°C. Although the net volume change is low, the axial expansion is highly anisotropic, with the *c* axis expanding and the *a* axis contracting in response to heat. This means, in practice, that in order to control intergranular stresses and prevent microcracking, the grain size of β -spodumene glass-ceramics should be held below 5 μm .

β -Spodumene glass-ceramics are fabricated via both bulk and powder sintering technique and have found wide utilization as architectural sheet, cookware, benchtops, hot-plate tops, heat exchangers, valve parts, ball bearings, and sealing rings. More recently, a β -spodumene ferrule for optical connectors was developed by Nippon Electric Glass by redrawing a cerammed preform of $\text{Li}_2\text{O}\text{-}\text{Al}_2\text{O}_3\text{-SiO}_2$ glass-ceramic. The cerammed preform contains about 50 wt% β -spodumene solid solution and a

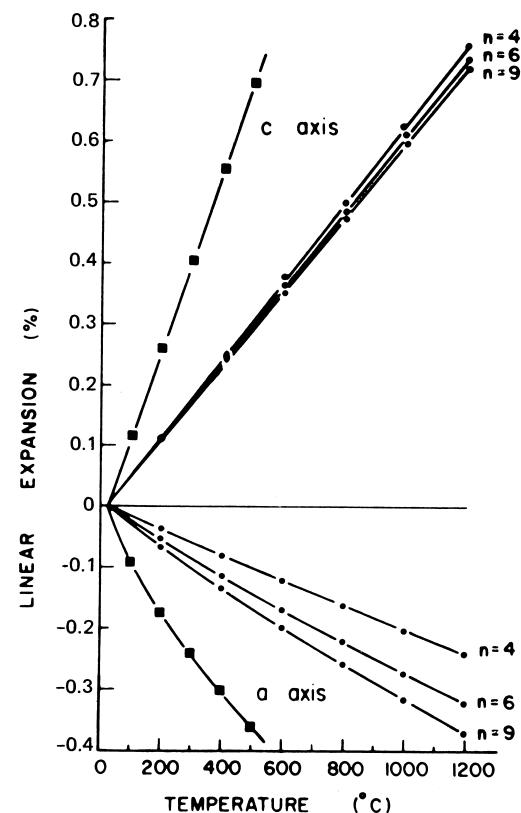


FIGURE 7 Linear thermal expansion of *a* and *c* axes; β -spodumenes: ●, keatite: ■. [After Ostertag, W., Fisher, G. R., and Williams, J. P. (1968), *J. Am. Ceram. Soc.* **51**(11), 651.]

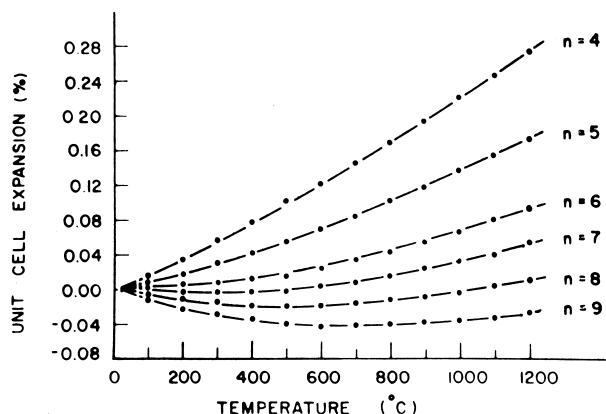


FIGURE 8 Volume thermal expansion in $\text{Li}_2\text{O}-\text{Al}_2\text{O}_3-n\text{SiO}_2$. [After Ostertag, W., Fisher, G. R., and Williams, J. P. (1968), *J. Am. Ceram. Soc.* **51**(11), 651.]

residual glass matrix. The preform is redrawn into microcapillaries with submicrometer accuracy, with no further crystallization occurring during the redraw. The microcapillaries are then processed to the near-net shape of the ferrule. The glass-ceramic ferrule exhibits excellent scratch resistance and good chemical durability, comparing favorably with conventional zirconia ferrules.

3. Cordierite Glass-Ceramics

Other low-expansion glass-ceramics, in the $\text{MgO}-\text{Al}_2\text{O}_3-\text{SiO}_2$ system, are based on the cordierite phase. These combine high strength and good thermal stability and shock resistance with excellent dielectric properties at microwave frequencies. These materials are made by bulk and powder sintering methods and are used in missile radomes and as high-performance multilayer electronic packaging.

4. Glass-Ceramics Based on Nonsilicate Crystals

a. Oxides. Glass-ceramics consisting of various oxide crystals in a matrix of siliceous residual glass offer properties that are not available with more common silicate crystals. In particular, glass-ceramics based on spinels and perovskites can be quite refractory and can yield useful optical, mechanical, and electrical properties.

Glass-ceramics based on spinel compositions ranging from gahnite (ZnAl_2O_4) toward spinel (MgAl_2O_4) can be crystallized using ZrO_2 and/or TiO_2 as nucleating agents. These glass-ceramics can be made highly transparent, with spinel crystals on the order of 10–50 nm in size (Fig. 9). The phase assemblage consists of spinel solid-solution crystals dispersed throughout a continuous siliceous glass. Possible applications for transparent spinel glass-ceramics include solar collector panels, liquid-crystal display screens, and high-temperature

lamp envelopes. Recently, nonalkali, nanocrystalline glass-ceramics based on Mg-rich spinel and enstatite (a chain silicate) crystals were developed for potential use as magnetic disk substrates in computer hard drives. These glass-ceramics possess elastic modulus over 140 GPa and, with grain sizes of less than 100 nm (Fig. 10), can be polished to an average roughness of 0.5–1.0 nm.

Glass-ceramics based on perovskite crystals are characterized by their unusual dielectric and electrooptic properties. Examples include highly crystalline niobate glass-ceramics that exhibit nonlinear optical properties, as well as titanate, niobate, and tantalate glass-ceramics with very high dielectric constants.

b. Phosphates. Many phosphates claim unique material advantages over silicates that make them worth the higher material costs for certain applications. Glass-ceramics containing the calcium orthophosphate apatite, for example, have demonstrated good biocompatibility and, in many cases, even bioactivity, making them useful as bone implants and prostheses. Mixed phosphate-silicate glass-ceramics based on fluorapatite and wollastonite, fabricated using conventional powder processing techniques, are bioactive and have good flexural strength of up to 200 MPa. The aforementioned combination of fluorapatite with phlogopite mica provides bioactivity as well as machinability. Cast glass-ceramics with microstructures comprising interlocking needles of fluorapatite and mullite provide high fracture toughness, with K_{1c} values greater than 3 MPa/m.

Certain glasses in the $\text{B}_2\text{O}_3-\text{P}_2\text{O}_5-\text{SiO}_2$ system melted under reducing conditions can yield a unique microfoamed material upon heat treatment. These materials

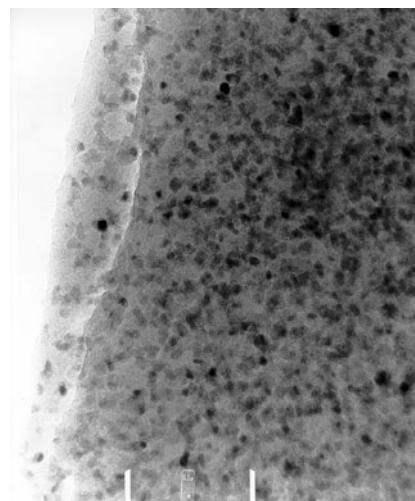


FIGURE 9 STEM micrograph of microstructure of transparent spinel glass-ceramic. Scale bar = 0.1 μm . [After Pinckney, L. R. (1999), *J. Non-Cryst. Solids* **255**, 171–177.]

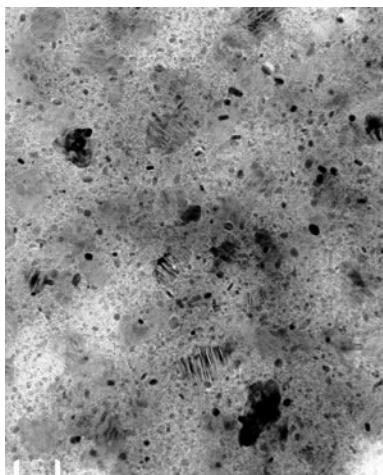


FIGURE 10 STEM micrograph of microstructure of high-modulus spinel/enstatite glass-ceramic. Scale bar = 0.1 μm . [After Pinckney, L. R., and Beall, G. H. (1997), *J. Non-Cryst. Solids.* **219**, 219–227.]

consist of a matrix of BPO_4 glass-ceramic filled with uniformly dispersed, isolated 1- to 10-mm hydrogen-filled bubbles. The hydrogen evolves on ceramming, most likely due to a redox reaction involving phosphite and hydroxyl ions. These materials, with DC resistivity of $10^{16} \Omega \text{ cm}$ at 250°C, dielectric constants as low as 2 and densities as low as 0.5 g/cm³, have potential application in electronic packaging.

More recently, a novel class of microporous glass-ceramics composed of skeletons of two types of titanium phosphate crystals has been prepared by chemical etching methods analogous to those used for Vycor (Corning) glasses. These materials offer promise for applications including oxygen and humidity sensors, immobilized enzyme supports, and bacteriostatic materials. Related lithium ion-conductive glass-ceramics, based on $\text{LiTi}_2(\text{PO}_4)_3$ solid solutions in which Ti^{4+} ions are partially replaced by Al^{3+} and Ga^{3+} ions, can be used for electric cells (solid electrolytes) and gas sensors.

c. Fluorides and Borates. Transparent oxyfluoride glass-ceramics, consisting of fluoride nanocrystals dispersed throughout a continuous silicate glass, have been shown to combine the optical advantages of rare earth-doped fluoride crystals with the ease of forming and handling of conventional oxide glasses. Fluorescence and lifetime measurements indicate that these materials can be superior to fluoride glasses both for Er^{3+} optical amplifiers because of greater width and gain flatness of the 1530-nm emission band and for 1300-nm Pr^{3+} amplifiers because of their higher quantum efficiency.

Transparent glass-ceramics based on submicrometer, spherical crystallites of $\beta\text{-BaB}_2\text{O}_4$ (BBO) in a glass of the

same composition demonstrated second harmonic generation to ultraviolet wavelengths; such materials offer promise as potential optical components. Other aluminoborate glass-ceramics may be useful as low-expansion sealing frits.

IV. CONCLUSIONS

In addition to continued growth in the areas of biomaterials and of zero-expansion, transparent glass-ceramics, future applications for glass-ceramics are likely to capitalize on designed-in, highly specialized properties for the transmission, display, and storage of information. The potential for glass-ceramics to play a key role in active and passive photonic devices for the rapidly growing telecommunications field is particularly exciting. Such products may include tunable fiber lasers and optical amplifiers as well as substrates for compensating for the temperature-induced wavelength variations in fiber Bragg gratings, versatile passive devices that can be used in a variety of optical components such as narrow-band add/drop filters, multiplexers, dispersion compensation, and amplifier module components.

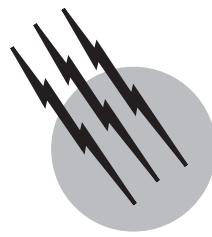
Whatever the future research trends in glass-ceramics are, it is certain that these materials, which combine the unique and diverse properties of crystals with the rapid forming techniques and product uniformity associated with glass production, will continue to proliferate as technology expands.

SEE ALSO THE FOLLOWING ARTICLES

CERAMICS, CHEMICAL PROCESSING OF • GLASS • PHOTOCHEMICAL GLASSES • PRECIPITATION REACTIONS

BIBLIOGRAPHY

- Beall, G. H., and Duke, D. A. (1969). "Transparent glass-ceramics," *J. Mater. Sci.* **4**, 340–352.
- Beall, G. H., and Pinckney, L. R. (1999). "Nanophase glass-ceramics," *J. Am. Ceram. Soc.* **82**, 5–16.
- Beall, G. H. (1992). "Design and properties of glass-ceramics," *Annu. Rev. Mater. Sci.* **22**, 91–119.
- Petzoldt, J., and Pannhorst, W. (1991). "Chemistry and structure of glass-ceramic materials for high precision optical applications," *J. Non-Cryst. Solids.* **129**, 191–198.
- McMillan, P. W. (1979). "Glass-Ceramics," 2nd ed., Academic Press, London, UK.
- Strnad, Z. (1986). "Glass-Ceramic Materials, Glass Science and Technology, Vol. 8," Elsevier, Amsterdam, The Netherlands.
- McHale, A. E. (1992). "Engineering properties of glass-ceramics," In "Engineered Materials Handbook, Vol. 4: Ceramics and Glasses," pp. 870–878, ASM International.



Metal Matrix Composites

Franck A. Girot

Azar P. Majidi

Tsu Wei Chou

University of Delaware

- I. Components
- II. Manufacturing Process
- III. Fiber–Matrix Interface
- IV. Mechanical Properties
- V. Structure–Performance Maps

GLOSSARY

Rheocasting Casting technique in which vigorous agitation of a semisolid alloy (temperature is between liquidus and solidus) before casting prevents the formation of dendritic properties in the primary phase.

Squeeze casting Casting technique in which liquid metal is pressed into a ceramic preform. This process provides close tolerance dimensions and high integrity.

Whiskers Short, single-crystalline fibers with the axis of growth in a crystallographic direction. They are 0.1–1 μm in diameter and up to several hundred micrometers long.

THE LARGE AMOUNT of research done on metals (e.g., aluminum, magnesium, copper, lithium, titanium, nickel, and chromium) since the 1970s has resulted in the development of materials with high specific properties, good thermal or electrical conductivities, and good corrosion resistance at low production costs. However, the field of application of these materials is restricted by their

low stiffness, the rapid drop of their properties at high temperature, and their poor fatigue resistance.

It has been well established in the past several years that the incorporation of ceramic fibers or whiskers into metal alloys is a very efficient way to improve their mechanical properties, particularly at high temperatures (above 300°C). Furthermore, these metal matrix composites (MMC) present different advantages with respect to organic matrix composites (where the matrix is a thermoset or a thermoplastic resin): they have higher transverse properties and can be used at elevated temperature. They also show better rupture energies, good thermal or electrical conductivities, good gas tightness, and insensitivity to water or ultraviolet ray emission.

The field of application of metal matrix composites has extended to aerospace, automotive, and electronic industries. The number of these applications is increasing as better processing conditions give rise to cheaper and higher performance materials.

The purpose of this contribution is first to describe the different components used in the manufacturing of metal matrix composites, then to briefly review the most widely

used manufacturing techniques now available, and finally to present the different properties we can expect from these materials.

I. COMPONENTS

The resulting properties of MMCs depend on (1) the ability of the matrix to support the fibers and to transfer the applied load to them and (2) the capacity of the reinforcement to sustain very high loads. The improvement of the properties of fibers and matrices is a main factor although the control of the interface and the choice of the different components are also important.

A. Matrices

The nature of the matrix will depend on the potential application of the composite. In that respect, one will choose aluminum or magnesium alloys for applications where low densities and high specific properties are required; nickel, titanium, or chromium for high-temperature applications; copper or aluminum when good electrical conductivities are necessary; copper or zinc when a low friction coefficient is important; and lead or lithium for batteries.

Before the fabrication of the composite, several difficulties will have to be overcome:

1. The nonwetting of most of the ceramic fibers by liquid metals and alloys can be solved by coating the fiber surface with TiB_2 for aluminum-, lead-, or copper-based alloys; K_2ZrF_6 for aluminum alloys; and Ni or Cu and SiO_2 for aluminum- or magnesium-based alloys. Alloying elements such as lithium or magnesium can also be added to the pure metal to improve the wetting.

2. The chemical reactions between ceramic fibers and the matrix can be minimized by coating the fibers with a protective layer, usually a ceramic material, by adding alloying elements to the matrix to obtain a system in equilibrium, or by reducing the contact time between fibers and liquid alloys.

3. The damage caused to the fibers can be minimized by optimizing the processing parameters.

B. Reinforcements

Because the fibers have to support high loads, they must have a high Young's modulus E and ultimate tensile strength (UTS). They must also have good fatigue properties, a low density, and a high-temperature capability. The only solids meeting these requirements are made of light elements with strong atomic bonds. So, it is not surprising that the different fibers available on the market are made

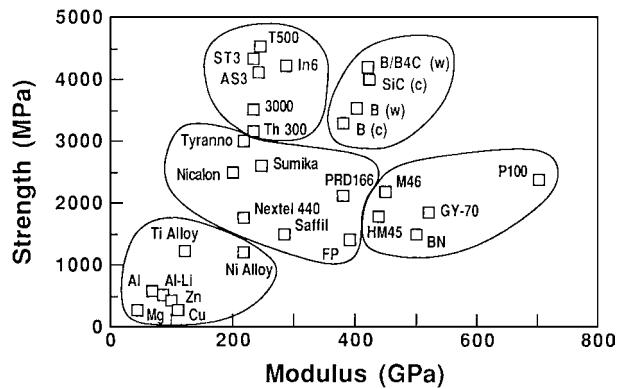


FIGURE 1 Fiber properties 1.

of elements such as boron, carbon, nitrogen, oxygen, aluminum, or silicon.

The different types of reinforcements available are in the forms of particles, fibers, and whiskers. The properties of various reinforcements are compared in Figs. 1 and 2. The potential applications of MMCs reinforced with particles are very limited, so only the last two types of reinforcements are presented.

1. Fibers Resulting from a Solid or Liquid Precursor

Carbon, boron nitride, silicon carbide, and alumina fibers have been obtained in this manner. The process consists of five steps:

1. Preparing and melting an organic precursor: polyacrylonitrile (PAN), usually for carbon fibers; boric anhydride for BN fibers; polycarbosilane for SiC fibers; or formacetate with alumina particles for alumina fibers.

2. The precursor is drawn out from the furnace (with viscosity around 1000 P) and reduced to a diameter of several micrometers.

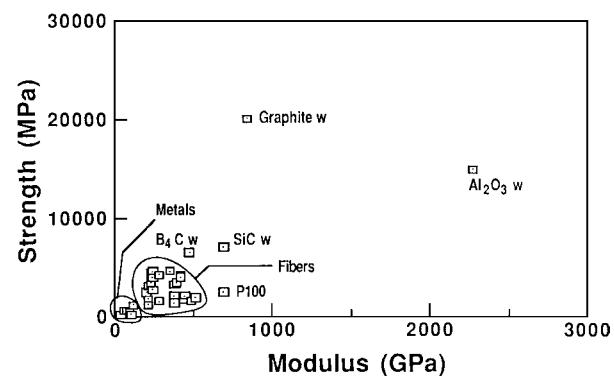


FIGURE 2 Fiber properties 2.

3. A stabilization treatment is then performed in order to make the fiber nonfusible (180–450°C depending on the fiber nature).

4. The fiber is pyrolyzed at a high temperature (1000–1400°C).

5. A very high temperature treatment (graphitization) follows the pyrolysis in the case of carbon (3000°C) and boron nitride (1800–2000°C).

After manufacturing the fibers are given a sizing treatment in order to protect them from damages and abrasion and to improve their compatibility with the different matrices they have to reinforce.

Two types of carbon fibers are currently available: high-modulus (HM) fibers with a modulus of ~340–700 GPa and UTS of ~1700–2400 MPa and high-strength (HS) fibers with a modulus of ~235 GPa and UTS of ~300–4500 MPa. These fibers can be easily woven into various shapes, but are very sensitive to oxidation and cannot be employed in oxidizing atmospheres above 600°C. In inert atmosphere they maintain their properties up to 1500°C.

Boron nitride fibers are of interest because of their electrical and thermal properties as well as their chemical stability at high temperatures. They can be used in air up to 800–1000°C.

Alumina and silicon carbide fibers have lower mechanical properties than carbon fibers but better chemical stability up to 1000–1200°C. They can be woven as carbon fibers but with some difficulties in the case of alumina fibers.

2. Fibers Resulting from a Gaseous Precursor

These types of reinforcement are essentially monofilaments of large diameter (100–200 μm) made of boron, boron carbide, or silicon carbide.

A fibrous substrate (a tungsten wire of 12 μm in diameter or a carbon wire of 30 μm) is run through a chemical vapor deposition (CVD) reactor, where gaseous reactants are introduced (hydrogen and boron trichloride or silane). The wire substrate is resistance heated to 1100–1400°C. Hydrogen and gaseous precursors then react in the reactor leading to the deposition of solid boron, boron carbide, or silicon carbide on the wire surface. After repeated depositions the filament reaches its final diameter and is then wound for use. The mechanical properties of these filaments are comparable to those of carbon fibers ($E \approx 400 \text{ GPa}$; $\text{UTS} \approx 3000–4000 \text{ MPa}$).

3. Whiskers

Whiskers are short single-crystalline fibers (0.1–1 μm in diameter and up to several hundred micrometers in

length) with the axis of growth in a crystallographic direction.

Their manufacturing process is based on usual processes of crystallogenesis with different mechanisms such as mass transport in gaseous phase, germination, and growth. Silicon carbide, silicon nitride, and alumina whiskers are manufactured by this process but at high costs.

Recently, a process that uses rice hules has been developed, which allows the manufacturing of silicon carbide whiskers at low price. It is well known that pyrolysis residues of some plants (e.g., rice hules) have a high silica content. These secondary products of the cereal industry are inexpensive and available in large quantities. The pyrolysis of rice hules leads to a mixture of organic compounds and silica. This mixture is then heated to high temperature, leading to the coking of the organic compounds, which transform into carbon. This carbon reduces the silica to give silicon carbide.

Whiskers are characterized by an elastic behavior up to rupture, and an ultimate tensile strength higher than that of continuous fibers, reaching the theoretical value ($\text{UTS} \approx E/10$).

II. MANUFACTURING PROCESS

The different techniques that are presented here are adapted for mass production (particularly the liquid- and semisolid-phase manufacturing processes) or for applications where a high purity of the different components is required (spray or powder metallurgy processes for electronics or aerospace).

A. Liquid-Phase Manufacturing Process

Liquid-phase fabrication methods are particularly suited for producing composite parts of complex shapes and the apparatus needed is relatively simple. Uniform fiber distributions can usually be achieved with little porosity in the matrix material. Several impregnation techniques have been used.

1. Vacuum Impregnation

Elements such as calcium or magnesium have been used to improve the wetting. Lithium is the most efficient element for the impregnation of alumina fibers. The alumina preforms are impregnated by aluminum-lithium (3%) alloy under vacuum in order to avoid the oxidation of lithium and to facilitate the infiltration by the liquid metal. Optimization of the process depends on the optimization of the following parameters: lithium content, preform temperature, and contact time between fiber and liquid metal.

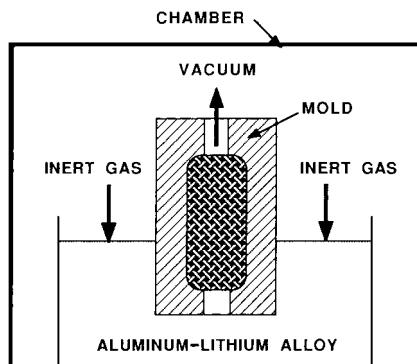


FIGURE 3 Vacuum impregnation.

The mechanical properties of these materials are generally high and close to the predicted values (Fig. 3).

2. Fluorides Treatment

A new casting process for the impregnation of fibrous ceramic preforms by aluminum alloys has been proposed in order to provide a low-cost technique for mass production of aluminum matrix composites. The method is based on a wetting enhancement treatment that involves the deposition of a fluoride K_2ZrF_6 from an aqueous solution at 90–100°C, within the pore network of the preform. The chemical reaction that occurs between liquid aluminum and the fluoride allows the dissolution of the alumina scale on the liquid-aluminum alloys. The alumina scale tends to prevent good wetting of most ceramic fibers. After such a treatment, a ceramic porous preform can be merely impregnated by gravity in a permanent mold heated at ~600°C, only by capillary effect (Fig. 4).

Although the process is easily performed, accurate experimental conditions (e.g., fluoride quantity and preform temperature) are necessary for a proper impregnation and good fiber-matrix bonding. Other types of fiber coatings (e.g., Ni, Si_3N_4 , and SiC) can be used along with the fluoride treatment to further control the interface properties.

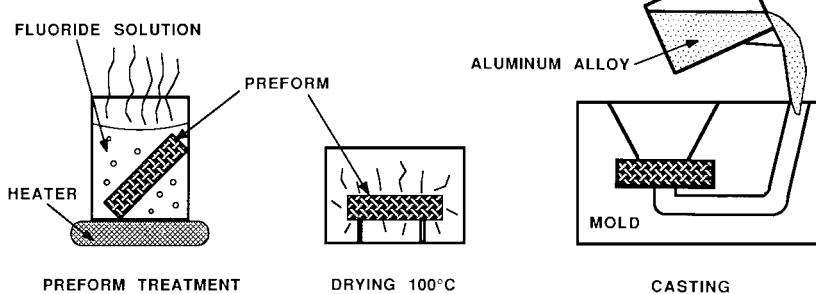


FIGURE 4 Fluoride treatment for wetting enhancement.

3. Squeeze Casting

Squeeze casting is a technique of liquid metal working under pressure that allows the manufacture of near net shapes with close tolerance dimensions and a high integrity. Usually, a preheated ceramic porous preform is placed in a die cavity, and a liquid metal is poured on. A piston applies the pressure on the liquid metal forcing it to impregnate the preform, as illustrated in Fig. 5. Two techniques can be distinguished: (1) direct squeeze casting (previously described) and (2) indirect squeeze casting where pressure enables to feed several molds.

Squeeze casting offers two major advantages:

1. Pressure compensates for the poor wetting of the ceramic preform with metal alloys (e.g., aluminum, copper, and zinc) regardless of the size of the reinforcement and the residual porosity of the preform before impregnation.
2. Solidification under pressure eliminates macro- and micropores leading to a better metalurgical quality of the composite.

The depths of impregnation depend on numerous parameters among which the most important are the preform and liquid-alloy temperatures, the physical characteristics and volume fraction of reinforcement, the contact angle between the ceramic and the liquid alloy, and obviously the pressure as well as the impregnation speed.

Actually, this process seems to be the most efficient because it allows the manufacture of finished parts at low cost and in a large scale. So, it is particularly adapted to the automotive industry for the manufacturing of connecting rods or pistons.

B. Semisolid-Phase Manufacturing Process: Compocasting

The compocasting process is the adaptation of the rheocasting technique to the manufacturing of discontinuous reinforced MMC. Rheocasting is a technique developed

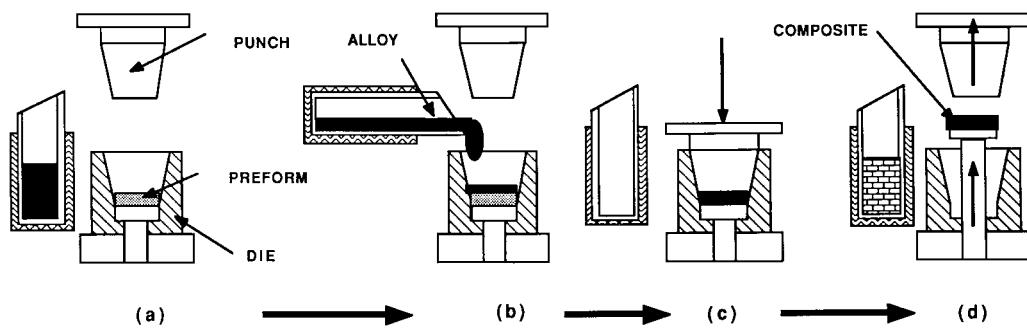


FIGURE 5 Squeeze casting is shown in steps (a)–(d).

by M. C. Flemings, which requires the vigorous agitation of a semisolid alloy (the temperature of the alloy is between the liquidus and the solidus) before casting so that the primary phase is nondendritic, giving a slurry with thixotropic properties and leading to a globular microstructure in the solidified alloy (Fig. 6).

MMC can be processed according to the following sequences:

1. The liquid-metal alloy is rapidly cooled down to temperatures in the liquidus–solidus range and is followed by a vigorous mixing, which prevents the primary phase to form dendrites and giving the mixture a low apparent viscosity.
2. At this steady-state temperature, the reinforcement is added to the melt and dispersed within the liquid by the solid primary phase.

3. Then the compound is cast when its fluidity is sufficient or a complementary technique such as injection or squeeze casting is used to form the composite when the viscosity is too high.

Usually, the resulting composites have a porosity content high enough to decrease significantly the mechanical properties. The complementary technique often allows to reduce the porosity content and to improve the metallurgical quality of the composite. However, it is possible to obtain cast composite parts with a low porosity content by performing the mixing under vacuum. Two variations of this process, namely mixing the reinforcement in a liquid matrix and reheating the compound to melt the matrix before casting, can overcome some of the drawbacks of this method but result in a more nonuniform fiber distribution.

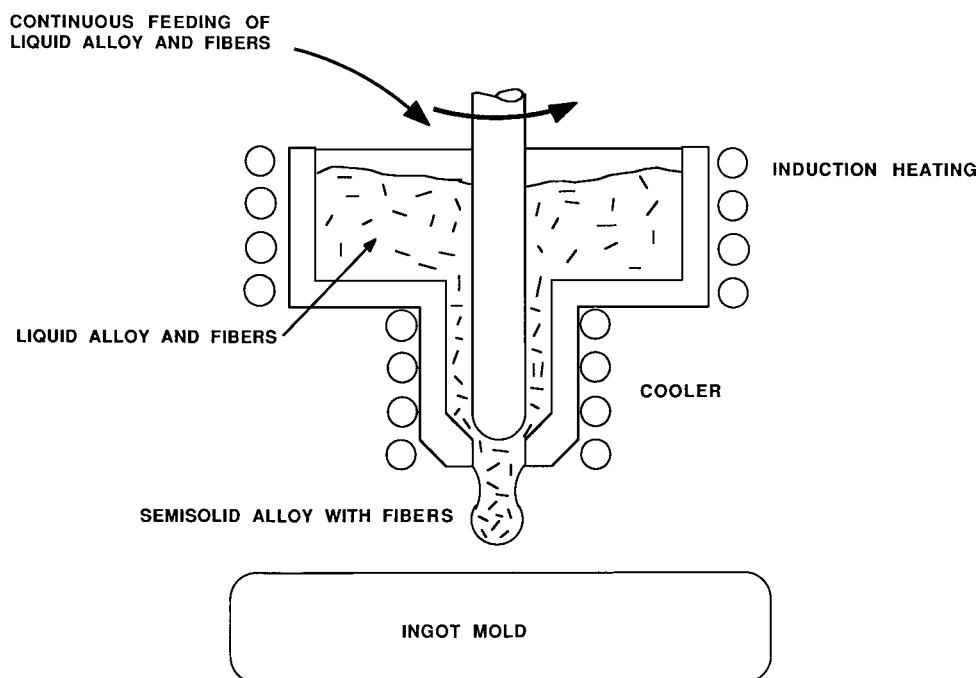


FIGURE 6 Continuous compocaster.

The compocasting process is particularly attractive for automotive applications (connecting rods), for friction application (bearings), and for electronics (heat sinks and alumina chip carrier supports).

C. Solid-Phase Manufacturing Process

1. Spraying and Diffusion Bonding

The spraying device has a hot zone produced by a gas flame, a gas explosion, or an electric arc through which a gas flows. Metal particles are fed into the gas flow, melted in the hot zone, and sprayed onto an assembly of fiber arrays. In the electric arc process, the arc is used to produce molten particles from the wire electrode, and the gas flow sprays them onto the fiber surface. A thin film of matrix is then formed, which binds to the fibers without physical or chemical damage to the fibers.

Diffusion bonding is then used to obtain highdensity composites. Preformed sheets are laid in predetermined orientations to achieve the fiber volume fraction and thickness required. Press forming achieves the bonding between the different sheets of fibers and matrix through the application of pressure and temperature.

Diffusion bonding can also be performed on alternative stacking of metal sheets and fiber weaves.

2. Powder Metallurgy

The incorporation of the reinforcement can be performed in the solid state by blending a powder of metal alloy and short fibers, whiskers, or particles. The reinforcing elements, whose volume fraction can be rather high, are generally well distributed. Pressure is then applied in order to form the mixture in near net shape part and to densify it before sintering. The part is heated to a high temperature, which depends on the nature of the metal alloy, and maintained at that temperature during sintering. Bonding is achieved usually by a diffusion phenomenon leading to composite with a good metallurgical quality.

However, the uniaxial or isostatic pressures, which are applied before sintering, are quite high so that some degradation of the reinforcement occurs, particularly when this sequence takes place at room temperature. This mechanical degradation can be reduced by applying pressure at high temperatures, eventually within the liquidus–solidus range. A schematic representation of MMC powder processing is given in Fig. 7. This process is less suitable when the reinforcement is long, but is particularly interesting in the case of reinforcement of small dimension such as whiskers. Moreover, secondary work can be applied to whisker-reinforced MMC by using processes such as hot rolling, extrusion, and drawing.

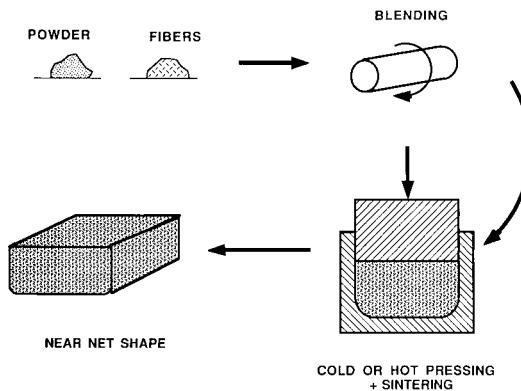


FIGURE 7 Metal matrix composites by powder metallurgy.

III. FIBER–MATRIX INTERFACE

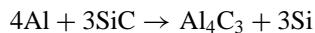
Metal matrix composites are generally nonequilibrium systems. Therefore, during fabrication or subsequent high-temperature exposure, diffusion of elements across the interface takes place, which often causes severe degradation of the fiber strength. Degradation is due to one or more of the following mechanisms; (1) formation of a brittle interfacial reaction layer; (2) reduction in fiber diameter; and (3) formation of voids in the fiber due to the rapid, outward diffusion of the elements. A thin reaction layer is often desirable as it promotes load transfer between the fiber and the matrix and thus increases the strength of the composite. A thick layer, on the other hand, is detrimental to the fiber strength because it fractures at a low stress and triggers fiber fracture. A fracture mechanics model for the strength of fibers with brittle reaction zones has been suggested. A summary of the interfacial characteristics of various metal matrix composites is given in the following.

A. Aluminum Matrix Composites

A severe reaction occurs between carbon fibers and aluminum alloys at temperatures above 550°C. Complete disappearance of the fibers may even take place at sufficiently high temperatures or after very long exposures. The primary reaction product is the brittle aluminum carbide, Al_4C_3 , phase which is often blocky, lath-shaped, or cuboidal. It has been found that pitch-based fibers are less prone to degradation than PAN-based fibers.

Interfacial reactions for alumina FP and SiC fibers in aluminum matrices are less severe. The reaction products in FP/Al–Li composites are predominantly LiAl_5O_8 spinel and LiAlO_2 and form at the grain boundaries in the fiber. Lithium is often added to the aluminum matrix to facilitate the wetting of the alumina fiber. For the system FP/Al–Cu (4.5%), research has shown no significant reaction.

The Nicalon SiC fiber is readily attacked by the molten aluminum to form Al_4C_3 by the following equation:



Although it may appear from this reaction that the presence of Si in the aluminum matrix will suppress the reaction, experiments with Al-Si (14.5%) alloy have shown that the opposite is true and the reaction zone actually thickens in the presence of Si in the matrix. The increased reaction in the Si-containing matrix is due to the presence of SiO_2 and free carbon in the Nicalon SiC fiber. In the case of CVD SiC monofilaments or some SiC whiskers where the reinforcement or some SiC whiskers where the reinforcement exhibits a higher degree of purity, the addition of silicon to aluminum may show the expected advantage. The beneficial effect of Si alloying on the stability of pure SiC-Al system has been demonstrated. In the case of SiC-whisker-reinforced aluminum-alloy composites, little reaction has been reported.

Boron filaments also react heavily with aluminum and aluminum alloys to form brittle interfacial zones, which severely lower the strength of the composite. The reaction layer is primarily aluminum diboride, AlB_2 , but may also contain other phases depending on the composition of the aluminum alloy. For example, in the case of B/6061 Al composites, a layer of AlB_{12} also forms on the fiber side in addition to the AlB_2 layer. A fiber fracture model for the thermally induced strength degradation of B-Al composites has been proposed. The model assumes a parabolic growth rate for the brittle interfacial reaction layer.

B. Magnesium Matrix Composites

Although magnesium alloys have a better wettability than aluminum alloys, there are still cases where fiber wetting is a limiting factor in the fabrication of magnesium-matrix composites. Katzman has recommended a silicon dioxide, SiO_2 , fiber coating from an organometallic precursor solution to promote fiber wetting by magnesium alloys. The coating was specifically developed for carbon fibers, but has also been successfully used to fabricate magnesium composites with FP alumina or Nicalon SiC fibers. In the case of high-modulus pitch-based carbon fibers, first a layer of amorphous carbon is deposited on the surface of the fiber to facilitate the adhesion of SiO_2 to the fiber.

In cast composites, the SiO_2 -coated carbon fiber reacts with the magnesium matrix to form an interfacial layer consisting of MgO and Mg_2Si . MgO is also the interface reaction product in FP-alumina-fiber-reinforced magnesium composites. Unlike the reaction layers in aluminum matrix composites, however, the presence of the MgO

reaction layer in magnesium matrix composites does not appear to seriously degrade the fiber strength.

No reaction has been observed in diffusion-bonded carbon fiber-magnesium composites.

C. Titanium Matrix Composites

Titanium and its alloys are highly reactive and attack most fibers. In boron-fiber-reinforced titanium composites, TiB_2 and TiB , reaction layers readily occur at the interface in the form of concentric rings with the TiB_2 ring being next to the fiber. Boron atom is smaller than titanium atom. Therefore, the outward diffusion of boron is faster than the inward diffusion of titanium, which causes void formation near the tungsten core in the filament. B_4C coating of the fiber provides a diffusion barrier which slows down the reaction. A SiC coating is much less effective as it is quickly destroyed due to the formation of titanium carbide and silicides. After longer annealing times Si diffuses into the matrix and Ti into the fiber, thus causing the formation of TiB_2 .

Titanium carbides and silicides are also the primary reaction products in composites of CVD-produced SiC monofilaments in Ti and Ti-6Al-4V matrices. The reaction consumes both the fiber and the matrix and, therefore, causes a severe reduction in the fiber diameter.

IV. MECHANICAL PROPERTIES

Fiber reinforcement significantly improves the strength and stiffness; resistance to fatigue, friction, and wear; and creep of the pure metals and alloys. The above improvements, however, are achieved at the cost of ductility and toughness, which result from the severe retardation of the plastic deformation in the matrix by the fibers.

The strength and stiffness in tension, compression, and shear of MMCs and the effect of thermal exposure on these properties have been extensively studied for aluminum, magnesium, and titanium matrix composites. These properties are often considerably below the predictions of the rule of mixtures. The reason is primarily fiber degradation during fabrication or subsequent thermal exposure as described above. The effects of fiber orientation and fiber spacing and diameter on the strength and stiffness of metal matrix composites have been reported.

J. L. Christian studied the fatigue strength of a variety of MMCs and found that the fatigue resistance of unidirectional Al matrix composites is higher in the 0° direction but lower in the 90° direction than those of Ti matrix composites. Heat treatment significantly improved the transverse fatigue strength but had little or

no effect on the axial fatigue resistance. In Ti-6Al-4V matrix reinforced with B, B(B₄C), and SiC filaments, J. M. Quenisset and coworkers found that a short thermal exposure at 850°C increased the resistance to fatigue crack propagation of composites reinforced with B and B(B₄C) fibers, while long exposures caused a severe drop in the strength and fatigue resistance. In SiC fiber composites, on the other hand, a short thermal exposure strongly decreased the fatigue resistance; however, the effect slowed down with time until, after 10 hr of exposure, all three composites exhibited a more or less similar degree of degradation. Several other investigations have also been conducted on the fatigue characteristics of MMCs under both mechanical and thermal cycling loadings. While the continuous fiber MMCs have shown a considerably higher axial fatigue resistance than their matrices, a reduction in the fatigue resistance has been reported in a randomly oriented short Saffil δ -aluminafiber-reinforced aluminum composites. Fiber fracture and a strong fiber/matrix bond were cited as possible reasons for the reduction.

The friction and wear properties of MMCs can be substantially better than those of the unreinforced matrix material. The wear rate has been found to be a strong function of the fiber orientation. The lowest rate is achieved when fibers are normal to the sliding surface. In carbon-fiberreinforced metal matrix composites, the type of the fiber has a significant effect on the wear rate. The low-modulus PAN-based fibers give the lowest wear and friction rates, while the high-modulus PAN-, rayon-, and pitch-based fibers give significantly higher rates of wear and friction.

The greatest benefit of fiber reinforcement of metals is to reduce their creep rate, thus increasing their maximum use temperature. Mechanistic models of creep of MMCs have been developed by A. Kelly and K. N. Street, M. McLean, and M. Lilholt. These models treat both cases of rigid fibers in a creeping matrix, characterized by ceramic fibers such as SiC or Al₂O₃ in metallic matrix, and creeping fibers in a creeping matrix, represented by MMCs reinforced by refractory metal wires. In a composite containing perfectly aligned, continuous fibers, the creep rate progressively falls until it reaches zero at a certain strain level. Short-fiber composites and continuous-fiber composites under off-axis loading, however, both exhibit a steady-state creep rate.

In addition to the composites based on unidirectional lamina or short fibers, multiaxially reinforced textile MMCs have also been developed and characterized. A comprehensive investigation of the strength, toughness, impact resistance, and friction and wear properties of three-dimensional braided FP alumina-fiber-reinforced

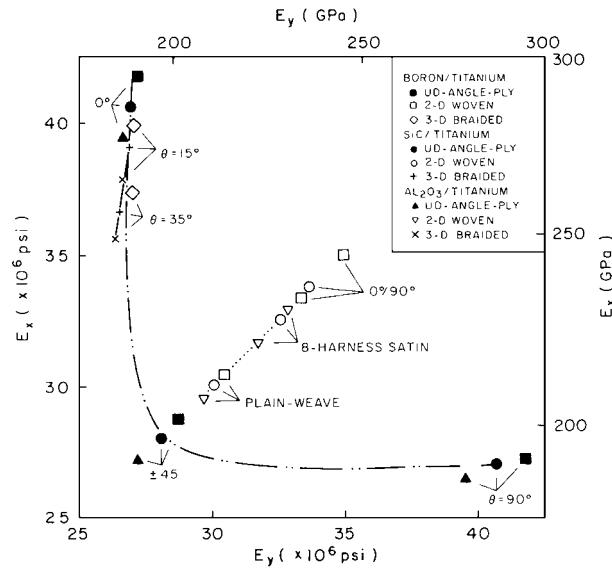


FIGURE 8 Structure–performance map 1.

Al-Li composites has been conducted by A. P. Majidi and coworkers.

V. STRUCTURE–PERFORMANCE MAPS

Considerable effort has been devoted by researchers to study the thermoelastic behavior of metal matrix composites and to evaluate the effectiveness of various reinforcement concepts. However, the analysis and experimentation performed on advanced composites are usually reported for individual systems; it is thus difficult

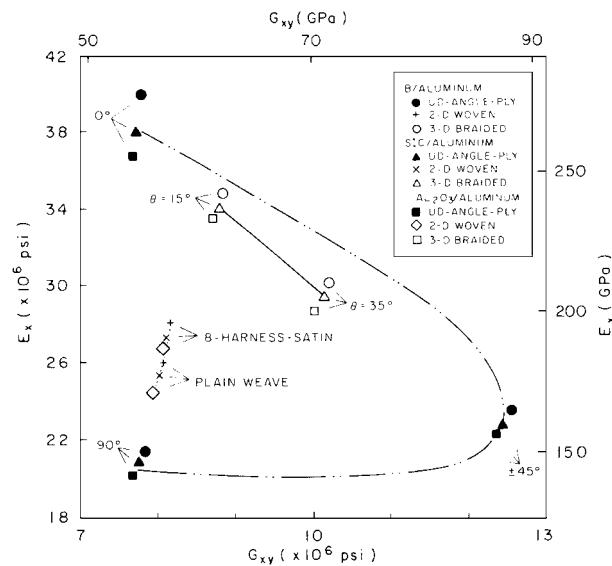


FIGURE 9 Structure–performance map 2.

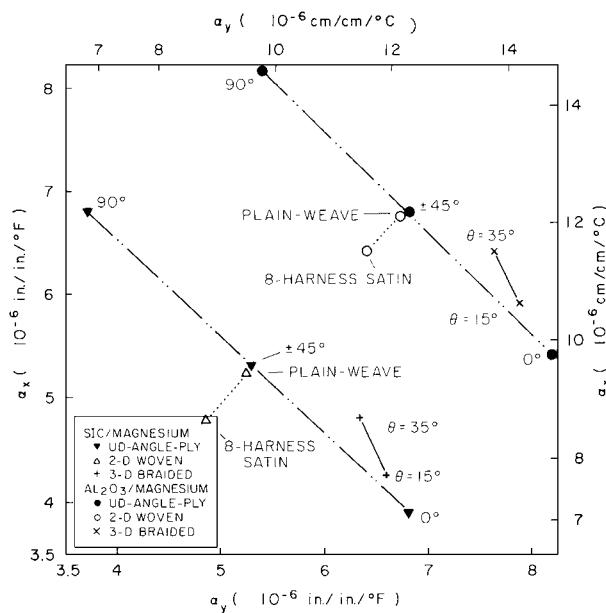


FIGURE 10 Structure–performance map 3.

to acquire a more comprehensive view. In this section, the results of extensive studies in the modeling of thermoelastic behavior of unidirectional laminated composites, as well as two-dimensional and three-dimensional textile structural composites, are integrated in the form of structure–performance maps.

The geometric configurations of the fiber reinforcements consist of angle-ply ($+θ/-θ$) and crossply ($0°/90°$) laminates of unidirectional laminae, two-dimensional fabric materials with increasing fabric repeating pattern size ranging from plain weave to eight-harness satin weave, and three-dimensional braided preforms.

The matrix metals included in this study are aluminum, titanium, magnesium, and a superalloy (MAR-M-200 of Martin Marietta Co.). The fiber materials are B, SiC, Al_2O_3 , and refractory wires of tungsten alloy (W-2 pct ThO_2) and molybdenum alloy (TZM). Figures 8–10 show the performance of the composite systems in terms of the Young's modulus (E_x and E_y in the two mutually orthogonal directions), shear modulus (G_{xy}), and thermal expansion coefficient ($α_x$ and $α_y$ in the two mutually orthogonal directions). In the braided composites the braiding angle, $θ$, is defined as the angle between the braiding axis and the average orientation of the fiber segments in the preform.

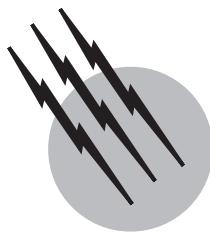
Through the construction of structure–performance maps, the relative effectiveness and uniqueness of various reinforcement concepts can be assessed. These maps should guide engineers in material selection for structural design and researchers in identifying the needs for future work.

SEE ALSO THE FOLLOWING ARTICLES

ALUMINUM • CERAMICS • COMPOSITE MATERIALS • FRACTURE AND FATIGUE • METALLURGY, MECHANICAL • POLYMERS, MECHANICAL BEHAVIOR

BIBLIOGRAPHY

- Chou, T. W., and Kelly, A. (1985). *Composites* **16**, 187–206.
- Chou, T. W., and Yang, J. M. (1986). *Metall. Trans. A* **17**, 1547.
- Clyne, T. W., and Withers, P. J. (1995). “An Introduction to Metal Matrix Composites,” Cambridge Univ. Press, Cambridge, MA.
- Clyne, T. W., and Simancik, F. (Eds.). (2000). “EUROMAT 99, Vol. 5, Metal Matrix Composites and Metallic Foams,” Wiley, New York.
- Delmonte, J. (1981). “Technology of Carbon and Graphite Fiber Composites,” pp. 41–87. Van Nostrand–Reinhold, Princeton, NJ.
- Girot, F. A., Fedou, R., Quenisset, J. M., and Naslain, R. (1987). “Proceedings of the American Society for Composites, Second Technical Conference,” Technomic, Lancaster, PA.
- Hayashi, T. (Ed.) (1982). “Progress in Science and Engineering of Composites,” Proc. ICCM 4, Tokyo, pp. 113–120, 1297–1305, 1315–1322, 1465–1472, Japan Soc. Composite Materials, Tokyo.
- Karandikar, P. G., and Chou, T. W. (1990). Manufacturing of C/Al composites by gravity infiltration and their characterization. In “Fundamental Relationships Between Microstructure and Mechanical Properties of Metal Matrix Composites,” (M. N. Gungor *et al.*, Eds.), pp. 59–69, TMS, Warrendale, PA.
- Karandikar, P. G., and Chou, T. W. (1991). Characterization of aluminum-matrix composites made by compocasting and its variations. *J. Mater. Sci.*
- Majidi, A. P., and Chou, T. W. (1987). Structure-reliability studies of three-dimensionally braided metal matrix composite. In “Proceedings of Sixth International and Second European Conference on Composite Materials,” Elsevier Applied Science, London.
- Majidi, A. P., Yang, J. M., Pipes, R. B., and Chou, T. W. (1985). Mechanical behavior of three-dimensionally woven fiber composites. In “Proceedings of the Fifth International Conference on Composite Materials, ICCM-IV” (W. C. Harrigan, J. Strife, and A. K. Dhingra, Eds.). The Metallurgical Society of AIME, Warrendale, PA.
- Martineau, P., Pailler, R., Lahaye, M., and Naslain, R. (1984). *J. Mat. Sci.* **19**, 2749–2770.
- Mileiko, S. T. (1997). “Metal and Ceramic Based Composites,” Elsevier, New York.
- Shehata, M. T., Leduc, T. R., Lemay, I., and Louthan, M. R. (1995). “Metallographic Techniques and Characterization of Composites, Stainless Steels, and Other Engineering Materials,” ASM International, Washington, DC.
- Shetty, H. R., and Chou, T. W. (1985). *Metall. Trans. A* **16**, 853–864.
- Terry, B., and Jones, G. (1990). “Metal Matrix Composites: Current Developments and Future Trends in Industrial Research Applications,” Elsevier, New York.
- Viala, J. C., Bosselet, F., Fortier, P., and Bouix, P. (1987). Chemical interaction between silicon carbide nicalon fibers and liquid aluminum or aluminum-silicon alloys. In “Proceedings of the Sixth International and Second European Conference on Composite Materials, ICCM IV & ECC 2” (F. L. Mathews, N. C. R. Buskell, J. M. Hodgkinson, and J. Morton, Eds.), Vol. 2, Elsevier Applied Science, London.
- Voyadjis, G. Z., and Kattan, P. I. (1999). “Advances in Damage Mechanics: Metals and Metal Matrix Composites,” Elsevier, New York.



Microporous Materials: Zeolites, Clays, and Aluminophosphates

Larry Kevan

University of Houston

- I. Zeolite Structure and Nomenclature
- II. Zeolite Synthesis
- III. Applications
- IV. Smectite Clays
- V. Aluminophosphates

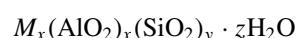
GLOSSARY

- AlPO** Abbreviation for crystalline, microporous aluminophosphates.
- Mesoporous** Porous materials with channels or cages of 20 Å to about 500 Å in diameter.
- Microporous** Porous materials with channels or cages from about 2–20 Å in diameter.
- SAPO** Abbreviation for crystalline, microporous silicoaluminophosphates.
- Smectite** A class of clays that swell by absorption of polar solvents.
- Zeolite** A crystalline, microporous aluminosilicate.

MICROPOROUS MATERIALS are typically defined as materials with high surface areas of several hundred $\text{m}^2 \text{ g}^{-1}$ containing cages and channels less than about 20 Å in diameter. Materials with larger channels and cages are called mesoporous. The most common and useful microporous oxide materials are zeolites, smectite

clays, and aluminophosphates. Since these materials have molecular-sized cages and channels they are also called molecular sieves. Typical applications include sorbents, ion-exchangers, catalysts, high surface area supports, and hosts for charge separation.

Zeolites are microporous crystalline aluminosilicates with cages and channels of molecular size ranging from about 2–15 Å in diameter. The framework typically consists of tetrahedrally coordinated aluminum and silicon based on AlO_4^- and SiO_4 units. The charge on AlO_4^- shows that the framework is negatively charged. This framework charge is balanced by cations which are typically exchangeable alkali ions and lead to important ion exchange properties for zeolites. In addition, water molecules are typically present within the zeolite structure and are readily lost and gained leading to important sorptive properties. A general formula for hydrated zeolites is as follows where M signifies a monovalent metal ion. A variety of zeolite



structures are found in nature and even more structures can be synthesized in the laboratory.

Clays are finely divided ($<2\text{ }\mu\text{m}$ diameter) crystalline silicates classified as phyllosilicates. These are often layered materials where a layer is composed of silicon-oxygen tetrahedral sheets and aluminum-oxygen or magnesium-hydroxyl octahedral sheets. Clays are designated as 1:1 layer materials when a layer contains one tetrahedral sheet and one octahedral sheet with oxygens shared between the two sheets. Also, 2:1 layer materials have layers composed of two tetrahedral sheets bordering an octahedral sheet. Among the 2:1 layer materials are the smectites. In smectites tetravalent silicon in a tetrahedral sheet may be partially replaced by trivalent aluminum and trivalent aluminum in an octahedral sheet may be partially replaced by a divalent ion such as magnesium. Such substitutions generate a negatively charged layer which is compensated by exchangeable cations within the interlayer space. Smectites have the characteristic of swelling or expanding by absorption of water or other polar solvents into the interlayer space which solvate the exchangeable cations. A structurally similar class of 2:1 layer materials which do not show swelling with solvent are the illites. Thus, smectites are also called “swelling clays” and can be commercially important as ion-exchangers and as catalysts. These applications are analogous to those for zeolites. Smectites are widely found in nature and some have also been synthesized in the laboratory.

Microporous crystalline aluminophosphates (AlPO) were first synthesized at the Union Carbide laboratories in 1982. Unlike zeolites and smectite clays, microporous AlPOs do not seem to be found in nature. Their synthesis seems to require organic additives such as organic amines or quaternary ammonium salts which may act as structure-directing agents. The charges on AlO_2^- and PO_2^+ show that the framework is neutral so no charge-compensating ions are required. By introducing a silica source into the synthesis mixture, Si often predominately substitutes for P to generate silicoaluminophosphates (SAPO) with a net negative framework charge that is balanced by cations. Thus, SAPO materials show ion exchange and catalytic properties. Both AlPO and SAPO structures include ones analogous to zeolite structures as well as unique structures not known among aluminosilicate zeolites.

I. ZEOLITE STRUCTURE AND NOMENCLATURE

Zeolite structures can be described or drawn in terms of their framework of tetrahedral atoms (Al and Si) denoted by points. These points are connected by bent O bridges

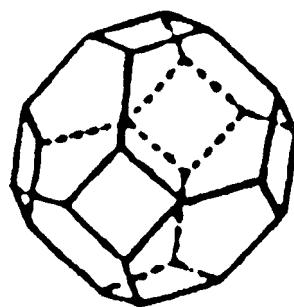


FIGURE 1 A truncated octahedron or cubed octahedron with eight hexagon faces and six square faces. In zeolite structures this is also known as a sodalite cage or β -cage.

denoted by straight lines so the straight line is a simplification. One such cage structure is a cubooctahedron or the truncated octahedron shown in Fig. 1. This is called a sodalite cage or β -cage and consists of eight 6-sided or 6-oxygen rings or 6-rings and six 4-rings. When these sodalite cages are connected by their 4-ring faces one has the zeolite called sodalite shown in Fig. 2. In this structure each sodalite cage is surrounded by six other sodalite cages. This is a fairly dense structure which is less typical than the more open structure of most zeolites. The largest 6-ring windows to the sodalite cage have about 2.2 \AA free access diameter for molecules to penetrate into this cage which has a free internal diameter of 6.6 \AA .

A. A, X, and Y Zeolite Structures

A more open structure is formed by connecting sodalite or β -cages on their 4-ring sides by cubes or double 4-rings. This structure is called zeolite A and is shown in Fig. 3. It is seen that a larger cage is formed in the middle of

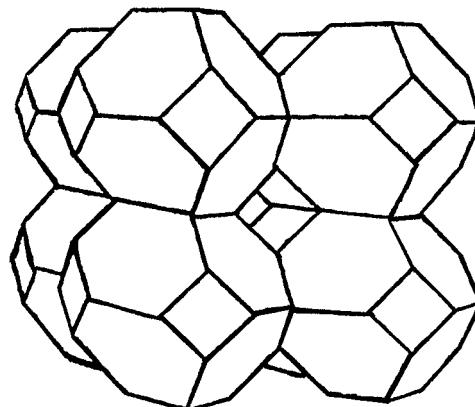


FIGURE 2 The sodalite zeolite structure made up of eight sodalite cages to form a space filling structure. Each point represents a tetrahedral Si or Al atom and each line represents an oxygen atom at its center.

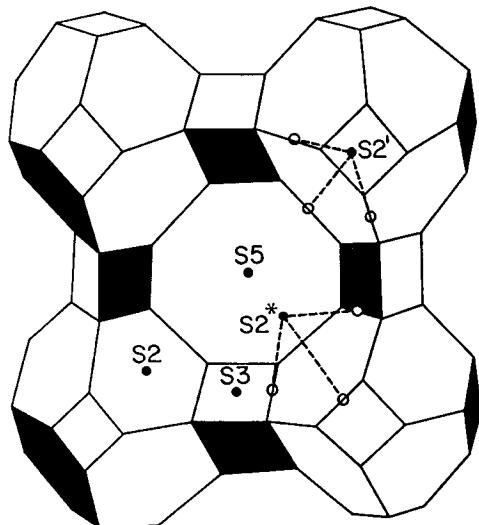


FIGURE 3 The zeolite A open structure composed of β -cages (sodalite cages) connected on their square faces by cubes. Probable cation sites are give by S5, etc.

the structure shown. This is called an α -cage which has a free internal diameter of about 11.4 Å with 8-rings as the largest openings which have about 4.1 Å free access diameter.

A still more open structure is formed by connecting sodalite cages on their 6-ring sides by hexagonal prisms or double 6-rings. This structure is called faujasite as shown in Fig. 4 which has two common synthetic forms with different Si/Al ratios. X zeolite has Si/Al \sim 1.3 and Y zeolite has Si/Al \sim 2.6. Again a large α -cage is formed which has a somewhat larger free internal diameter of

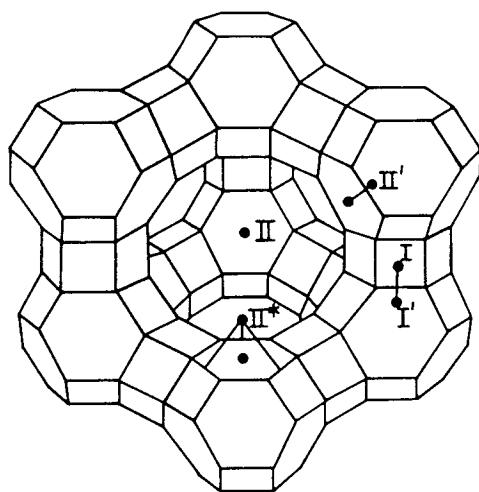


FIGURE 4 The zeolite X or Y open structure composed of β -cages connected on their hexagon faces by hexagonal prisms. Probable cation sites are indicated by I, etc.

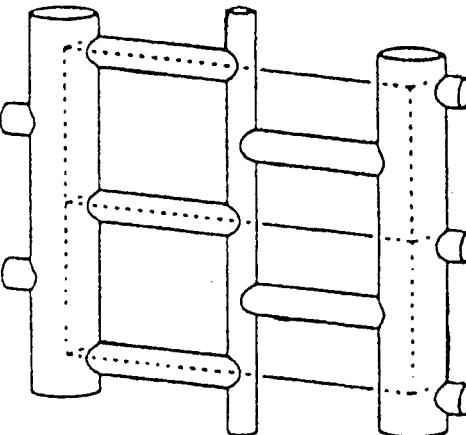


FIGURE 5 The mordenite zeolite structure represented in terms of one plane of channels. The large tubes are 12-ring channels and the small tubes are 8-ring channels. Note that there are both parallel and perpendicular 8-ring channels that indirectly connect the large 12-ring channels which are all parallel.

about 12.5 Å than does the α -cage in A zeolite. The largest openings to the faujasite α -cage are 12-rings with about 7.4 Å free access diameter.

A, X, and Y zeolites are cage-type structures with access between cages by smaller windows (n -rings) than the free cage diameter. A variety of other cage-type zeolite structures also exist.

B. Mordenite, ZSM-5, and ZSM-11 Structures

Other zeolite structures seem best described in terms of channels rather than cages. An important group of channel-type zeolites is based on 5-rings linked into a series of chains which form major channels. Mordenite is one example which has 12- and 8-ring channels. The channel structure of mordenite is represented in Fig. 5. Two interesting synthetic channel-type zeolites are ZSM-5 and ZSM-11 whose channel structures are shown in Fig. 6. These zeolites have only 10-ring channels but there are

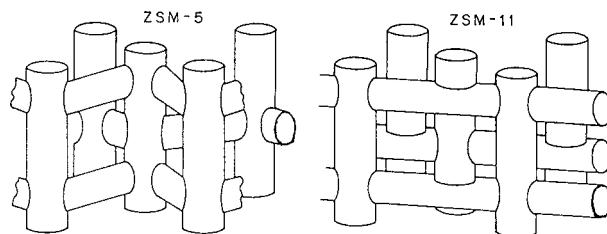


FIGURE 6 Channel structure of ZSM-5 and ZSM-11 zeolites. Both structures are composed of intersecting 10-ring channels. In ZSM-11 the two sets of channels are straight and nearly circular while in ZSM-5 one set of channels is straight while the other set is zig-zaged and both sets of channels are elliptical.

two sets of channels that intersect. In ZSM-11 the two sets of channels are both straight and are nearly circular ($5.3 \times 5.4 \text{ \AA}$). But in ZSM-5 straight and elliptical channels ($5.4 \times 5.6 \text{ \AA}$) intersect with zig-zag and elliptical channels ($5.1 \times 5.5 \text{ \AA}$).

C. Zeolite Nomenclature

The word “zeolite” is believed to originate from the Greek words *zein* (to boil) and *lithos* (stone) or boiling stone which refers to the visible loss of water when natural zeolites are heated. Zeolite was first used by the Swedish mineralogist Cronstedt in 1756 to describe a new mineral species with that characteristic.

Over 40 natural zeolites are known and have a wide variety of mineral names including chabazite, faujasite, heulandite, mordenite, and stilbite. Zeolite synthesis began in 1938 with the work of Barrer in England. Barrer synthesized several natural zeolites and also found some new phases. He designated these materials by the cation used in the synthesis and a capital letter like Li-A, Na-B, etc. From 1949 Union Carbide in the United States pioneered low-temperature hydrothermal synthesis of zeolites and also used capital letters to designate these materials like A, L, X, and Y. The letter designations used by Barrer and by Union Carbide unfortunately did not refer to the same structure and this is perhaps the origin of the diverse and nonsystematic nomenclature for synthetic zeolites today. In addition, Union Carbide and Mobil Oil also used Greek letters to designate some new synthetic zeolite phases. The most common of these is zeolite beta synthesized by Mobil.

Currently, new zeolite phases are designated by a two-to four-letter abbreviation for the company or institution where the synthesis was performed followed by a hyphen and number indicating a specific characterized phase. Thus, ZSM-5 refers to Zeolite Socony Mobil structure 5. ZSM-1 is not used because that preparation turned out not to be a new single phase material. So the structure numbers are not sequential, which can further add to the confusion. Other abbreviations indicate MCM for Mobil Composition of Matter, SSZ for Socal Silica Zeolites, DAF for Davy Faraday Laboratory, and VPI for Virginia Polytechnic Institute among others.

D. Zeolite Structure Codes

The structure of a zeolite is described by its chemical composition, its tetrahedral framework topology, and the location of exchangeable cations and water. To help systemize the many names used for natural and synthetic zeolites, the framework topology is assigned a unique three-letter code by the Structure Commission of the International

TABLE I Structure Codes of Commonly Studied Zeolites

Structure code	Name
ANA	Analcime
BEA	Beta
CAN	Cancrinite
CHA	Chabazite
ERI	Erionite
FAU	Faujasite or X or Y
FER	Ferrierite
HEU	Heulandite
LEV	Levyne
LTA	Linde Type A or A or ZK-4
LTL	Linde Type L or L
MAZ	Mazzite
MEL	ZSM-11
MFI	ZSM-5
OFF	Offretite
RHO	Rho
SOD	Sodalite
STI	Stilbite

Zeolite Association. This structure code is preferentially based on a previously discovered natural mineral, such as CHA for chabazite. There are a few exceptions when a zeolite was synthesized before being discovered in nature such as zeolite beta (BEA) for which the natural mineral is tschernichite. Table I shows the structure codes for some commonly studied zeolites. There are over 98 different structure codes and this number is slowly increasing. These include not only aluminosilicate zeolites but also microporous AlPOs, SAPOs, and silicates.

It is helpful to include the structure code in parenthesis after the abbreviated name of the zeolite such as ZSM-5 (MFI), but this is seldom done in the current literature. However, it is common to designate the exchangeable cation or cations followed by a hyphen before the zeolite name abbreviation such as Na-A, K-X, CsNa-X, H-Y, and Na-ZSM-5. The Si/Al ratio is also important to designate but is often left out. Of commonly studied zeolites A(LTA) has Si/Al = 1, ZK-4 (LTA) has Si/Al > 1, X(FAU) has Si/Al ~1.3, Y(FAU) has Si/Al ~2.6, mordenite (MOR) has Si/Al ~5, and ZSM-5 has Si/Al ~30 typically but is highly variable up to nominal infinity.

II. ZEOLITE SYNTHESIS

In nature zeolites are often found in volcanic rocks. This suggests their formation under geological conditions of high temperature and pressure. In the earliest work on

reliable zeolite synthesis by Barrer in the 1940s hydrothermal conditions of high temperature, autogenous pressure, and high salt concentrations were used to synthesize some naturally occurring zeolites. However, in the 1950s Milton and Breck at Union Carbide showed that zeolites can be successfully formed under mild hydrothermal conditions of about 100°C, atmospheric pressure, and high water concentrations. They reported the synthesis of zeolite A in 1956 which is considered to be the first synthesis of a well-characterized new zeolite unknown in nature.

Hydrothermal synthesis of zeolites involves preparation of a well-mixed aqueous inhomogeneous gel from a silica source and an alumina source under high pH of 10–13. Typical aluminum sources are sodium aluminate and aluminum alkoxides. Typical silicon sources are soluble silicates, silica sols, and fumed silicas. The predominant alkali or alkaline earth cation is also important to generate specific structures. In addition to these ingredients, the concentrations, order of mixing, degree of mixing, temperature, time, and sometimes pressure must be controlled. There are few general rules, but a large body of synthetic information has been developed for reliable synthesis of many zeolites. Since more than one zeolite phase can be produced from a given gel composition, the optimal crystallization time is sometimes critical and this can vary with pH and other factors.

Organic cations are also used to help generate certain structures. For example, tetrapropylammonium ion can be used to synthesize ZSM-5. Such ions are often called templates, but their exact role is unclear and they are better called structure-directing agents as are simple alkali cations.

The exact chemical natures of the precursor chemical species are unclear as is the mechanism of building up a specific crystal structure. Various silicate and aluminate species have been identified by nuclear magnetic resonance and by Raman spectroscopy, but detailed synthesis mechanisms are lacking. This situation is likely to continue since a given zeolite structure can typically be synthesized with a variety of structure-directing agents—organic and inorganic. Thus, there seem to be many pathways to arrive at a given zeolite product.

Since zeolites are crystalline, the primary characterization method is powder X-ray diffraction. The single crystals are generally too small to apply single crystal X-ray diffraction. Also, since zeolites are microporous, the surface area determined by nitrogen adsorption is critical for characterization of good material. Other structural characterization is done by a variety of spectroscopies. Infrared spectroscopy can give information about hydroxyl groups and acid sites. Nuclear magnetic resonance gives information about the environments of magnetic nuclei like

^{29}Si and ^{27}Al . Electron spin resonance gives information about the environments of paramagnetic cations. A variety of other physical techniques have also been used to study zeolites.

III. APPLICATIONS

Zeolites are applied as selective ion exchangers, as sorbents and as catalysts for a variety of chemical processes. Clinoptilolite also has an amusing application as a pig dietary supplement, particularly in Japan. A 5 wt.% supplement of natural clinoptilolite to pig feed generally enhances weight gains and protects from some diseases. This effect seems less noticeable for other animals but clinoptilolite supplements to the feed of beef cattle are used in Hungary. Clinoptilolite has generally been used because it is the most abundant natural zeolite and so is cheap. The actual chemical effect of clinoptilolite as an animal feed supplement is unclear, but it may involve ion exchange control of ammonium and other ions as well as of some blood proteins in the gut. It is also clear than zeolites, even in relatively large quantities, are nontoxic.

A. Selective Ion Exchangers

Selective ion exchange by zeolites is a sensitive function of zeolite structure, silicon to aluminum ratio, hydrated ion size, strength of ion hydration, and other factors. For example, the selectivity for ion exchange of alkali metal cations in zeolite Y with $\text{Si}/\text{Al} \sim 2.6$ is



while for zeolite X with $\text{Si}/\text{Al} \sim 1.3$ it is



The largest volume application of zeolites as ion exchangers is as laundry detergent “builders” to enhance the cleaning efficiency of laundry detergents by removing Ca^{2+} and, to a lesser extent, Mg^{2+} ions from “hard” water. In the 1970s synthetic zeolite A was found to be the most effective zeolite for this purpose and largely replaced polyphosphates as detergent builders since phosphates were linked to environmental damage. Na-A zeolite is effective for removing Ca^{2+} but is much less effective for removing Mg^{2+} . The wide use of zeolite A in laundry detergents has generated no deleterious environmental or human health effects so it seems compatible with the “green” chemistry focus of today.

Zeolites are also important as ion exchangers to treat liquid nuclear waste which contains radioactive ^{137}Cs and ^{90}Sr since the water for storage of spent nuclear fuel

elements has pH 11.4. Since zeolites are not generally structurally stable in acid solutions, they are effective for ion exchange applications in basic solutions. Synthetic chabazite (Linde AW-500), natural mordenite, and natural clinoptilolite have been used effectively for nuclear waste treatment.

Waste water treatment has also benefited from the use of natural clinoptilolite to remove ammonia and ammonium ions by selective ion exchange. Clinoptilolite has been effectively applied to treat wastes from the rapidly growing industry of fish farming. Clinoptilolite has also been used to remove Fe^{3+} from boiler waters and Zn^{2+} and Cu^{2+} from nonferrous metal production wastes.

Finally, clinoptilolite and mordenite are widely used in Japan for soil beneficiation to control soil pH, moisture content, and manure malodor. These effects are likely related to selective ion exchange which enhance the natural ion exchange properties of soil due to its clay mineral content. Soil applications are most economic where natural zeolites are abundant as in Japan. However, a similar application to control malodor in cat litter boxes in the United States is not constrained by such economic considerations.

B. Sorbents

A second major application area for zeolites is their use as sorbents. A and X zeolites have excellent affinity for water and are extensively used as drying agents in refrigeration systems, in power transformers filled with hydrocarbon liquids, and in various industrial hydrocarbon gas streams. A novel use as a static desiccant is zeolite incorporation between double-paned windows to prevent fogging. X zeolite is also used to adsorb undesirable sulfur- and nitrogen-containing molecules from gaseous systems.

The molecular-sized channels and cages in zeolites suggest their use as "molecular sieves." The molecular sieving action is complex since several size channels and cages are usually present. In addition, electrostatic aspects are very important in addition to geometry. However, numerous practical applications exist such as the separation of n-alkanes from iso-alkanes and the separation of C₈ aromatic isomers. Linde developed 3A, 4A, and 5A molecular sieves where the number indicates the approximate largest pore diameter in angstrom that will pass molecules of that approximate size. Zeolite 4A is Na-A and 3A (K-A) and 5A (Ca-A) are produced from it by ion exchange. The ion exchange is incomplete, however, so that 3A is actually K, Na-A and 5A is Ca, Na-A. The different sieving effects of these materials show that the cations are situated in the cage openings and affect the access of molecules into the zeolite. Since K⁺ is larger than Na⁺, it blocks the opening more and forms a material with a smaller net opening for

molecular access. The cation size effect becomes more complicated in a material like mordenite with both 8- and 12-ring channels of different size because a cation of a certain size may preferentially block the larger or the smaller channel and direct molecular access to the other channel.

X and Y zeolites with specific exchangeable cations are also used for a number of liquid-phase separation processes such as para-xylene from its isomers, ethylbenzene from its isomers, and fructose from sucrose. These applications are called Sorbex processes and were developed by the UOP Corporation. These processes depend on both geometrical molecular sieving and electrostatic interactions and it is not really possible to separate these two aspects.

C. Catalysts

The most profitable or value-added application of zeolites is as catalysts for catalytic cracking of crude oil and related processes. Zeolite X was first used as such a cracking catalyst in 1962 and since that time faujasite zeolites have become dominant in that area. This has resulted in a great deal of practical research on zeolite synthesis and modification for this purpose. Acid sites on the zeolite are the important sites for catalytic cracking. These are maximized by ammonium ion or polyvalent metal ion exchange followed by heating in air which is called calcination. Both Bronsted and Lewis acid sites are important and the specific preparation process is key to obtain the optimum acidity for a specific chemical process.

Shape selectivity of the reactant or product by the molecular sieving action of zeolites is also important for specific catalytic activity. As described above, this can be controlled by utilization of a zeolite with a particular largest n-ring size and by judicious cation exchange. Crystallite size also affects the diffusion rate within a zeolite and is likewise important for shape selectivity.

Zeolites are also suitable for the preparation of bifunctional catalysts which can simultaneously act as both hydrogenation and dehydrogenation catalysts. This is accomplished by supporting a dispersion of reduced metal on a zeolite to act as dehydrogenation sites while zeolitic acid sites act as hydrogenation sites. Ion exchange of a metal salt into a zeolite followed by hydrogen or carbon monoxide reduction at an appropriate temperature can generate small metal clusters that act as dehydrogenation sites. These metal clusters are formed both inside the zeolite framework and on its external surface.

Catalytic cracking of petroleum to breakdown long-chain hydrocarbons to desirable C₁–C₆ fractions used in gasoline typically uses a fluidized bed of zeolite catalyst consisting of rare-earth exchanged Y zeolite or

H-Y zeolite. The catalyst is cyclically regenerated by heating in air at 600–700°C.

Hydrocracking converts heavy gas oils to heating oil, jet fuel, diesel fuel, and petrochemical feedstocks. Bi-functional zeolite catalysts are typically used for hydrocracking and consist of H-Y, ultrastable Y (more thermally stable H-Y), and dealuminated Y zeolites with dispersed Co-Mo, Ni-W, and Pt-Pd metals.

Selectoforming involves selective cracking of n-alkanes to propane to eliminate low octane fractions and upgrade the fuel. Ni-erionite and Ni-clinoptilolite are used for this catalytic process. Hydroisomerization converts n-alkanes to branched alkanes with higher octane number and uses Pt-mordenite. Dewaxing is the selective hydrocracking of very long hydrocarbons to a liquid propane fraction and uses Pt-mordenite or H-ZSM-5. Other important processes catalyzed by ZSM-5 include benzene alkylation and xylene isomerization.

A final example of important processes catalyzed by zeolites is methanol to gasoline conversion. This is particularly important to countries with less access to oil. Natural gas or coal is steam reformed to generate syngas which is a mixture of CO and H₂. The Fischer-Tropsch process is used to convert syngas to methanol and ZSM-5 converts methanol to high octane gasoline.

IV. SMECTITE CLAYS

Smectite clays are swelling clays that contain an interlayer space which can expand by the absorption of a suitable solvent. Figure 7 shows a typical structure of smectite clays where only the oxide lattice is explicitly shown. Each clay layer is made up of three sheets consisting of two tetrahedral sheets sandwiching an octahedral sheet. In the center of each tetrahedron or octahedron are cations such as aluminum, silicon, magnesium, etc. The thickness of this clay layer is 0.96 nm and the interlayer space can range from zero to complete separation or delamination of the layers. For a well-ordered system that has not delaminated, X-ray diffraction peaks are observed corresponding to the distance (d_{001}) between the basal oxygen surfaces of two successive layers as shown in Fig. 7. Also shown in Fig. 7 is the interlayer space. A hexagonal cavity is indicated by the arrow which corresponds to a cation location when it interacts strongly with the basal oxygen surface of the interlayer space, such as when most solvent molecules have been removed.

Three common clay structures will be discussed as representative: montmorillonite, beidellite, and hectorite. The ideal unit cell formulas are shown below, where the first parenthesis represents the cation composition of the octahedral sheet and the second parenthesis represents the

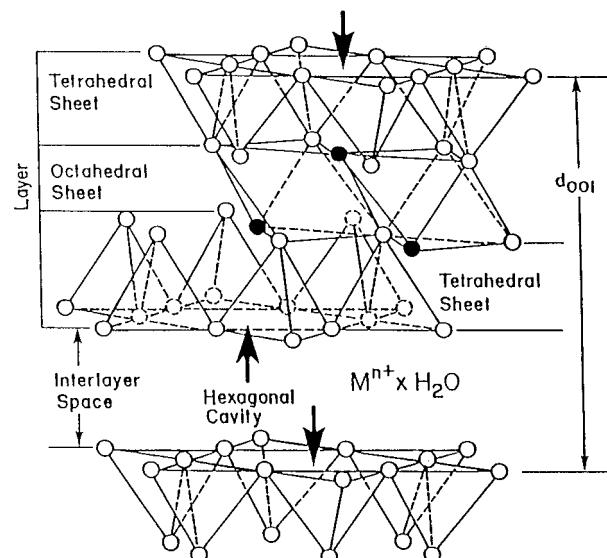
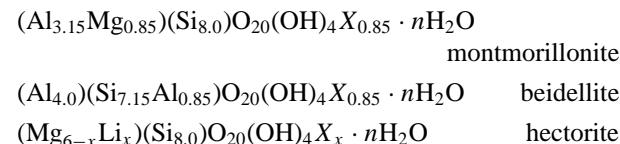


FIGURE 7 Smectite clay structure where only the oxide lattice is specifically shown indicating tetrahedral and octahedral arrangements of oxygens in the tetrahedral and octahedral sheets. Some oxygens are converted into hydroxyl groups as shown by the black circles. Cations such as silicon and aluminum, magnesium, iron, and lithium are located in the centers of the oxygen tetrahedrons and aluminum, magnesium, iron, and lithium are located in the centers of all or part of the oxygen octahedrons. The thickness of a clay layer is shown by d_{001} which is determined by X-ray powder diffraction. M is an exchangeable cation in the interlayer space which is typically solvated. A hexagonal cavity is indicated which corresponds to an exchangeable cation location when most of its solvent molecules have been removed.

cation composition of the two tetrahedral sheets in each layer.



X represents a monovalent exchangeable cation in the interlayer space. Note that both montmorillonite and beidellite have only four ions in the octahedral sheet. This means that some of the octahedra are vacant since there are six octahedral vacancies per unit formula. Thus montmorillonite and beidellite are called dioctahedral clays. Hectorite is a trioctahedral clay which contains six cations in its octahedral sheet and has all of the octahedra occupied by a cation. Laponite is a trade name of a commercially available synthetic hectorite. All three clays are found in nature. In addition, hectorite and beidellite have been synthesized in the laboratory.

Hectorite can be synthesized by preparing a gel of the proper composition in basic solution of about pH = 9 followed by refluxing for 7 days. Fluorohectorite can also

be synthesized in which the hydroxyl groups are replaced with fluoride. This is done by adding LiF to the gel before refluxing. It is also possible to incorporate transition metal ions into the hectorite framework in either octahedral or tetrahedral positions by incorporation of the appropriate metal ion salt into the synthesis gel. If CuCl₂ is added Cu(II) is found to substitute into the octahedral sheet.

Beidellite synthesis is more complex and requires high pressure. An appropriate gel is made in basic solution and then dried. The dried gel is placed in a gold tube and sealed. Gold seems required to maintain the purity of the gel to impurity metal ions. This synthesis bomb is heated at 330°C at autogenous pressure for 7 days to form beidellite.

Another synthetic variant of smectite clays is called pillaring. The solvent expandability of smectite clays limits their thermal stability. Their interlayer collapse by heating above ~200°C can be prevented by pillaring. Pillared clays are obtained by exchanging the charge-compensating cations in the interlayer space of smectite clays with metal-hydroxy polymeric cations such as Al₁₃-hydroxy and Zr₄-hydroxy polymers followed by thermal treatment which generates a robust pillar spanning the interlayer space. Pillaring with Al₁₃-hydroxy involves predominant Al₁₃O₄(OH)₂₈³⁺ while pillaring with Zr₄-hydroxy involves predominant Zr₄(OH)₁₄²⁺. These pillars have different sizes which is a way to control the height of the interlayer space in a pillared clay. This height can range from 5 to 20 Å. These pillared clays are thermally stable above 500°C and can be used as catalysts for a variety of organic reactions. The pillar density can be varied but the distribution of the pillars is usually nonuniform. Thus, few regular size channels or pores are formed as in zeolites. Since the upper size range for the pillars approaches 20 Å and the maximum channel size is larger than in zeolites, pillared clays offer the possibility of processing larger molecules than is possible with zeolites. This potential has not been realized because subsequent development has led to mesoporous silicas with channel sizes from 20 to ~300 Å.

V. ALUMINOPHOSPHATES

Aluminophosphates also form microporous molecular sieve materials. However, natural mineral AlPOs are dense phases that are not really microporous. For example, berlinitite is the AlPO structurally analogous to quartz and AlPO analogs of the dense silica phases of tridymite and cristobalite have been synthesized. Natural AlPO hydrates (variscite and metavariscite) and hydroxides (senegalite and wavelite) are also known.

In variscite the phosphorus is tetrahedral but the aluminum is octahedral due to coordination with the oxygens of two water molecules. Loss of water causes structural collapse. Various AlPO hydrates were synthesized in the 1960s including variscite, but many mixed phases resulted and structure characterization was incomplete.

True, well-characterized microporous AlPOs termed AlPO₄ or later simply AlPO materials were first synthesized in 1982 by a group at Union Carbide. Equivalent microporous materials have not yet been found in nature. The microporous AlPO materials are synthesized hydrothermally at 130–200°C at an acidic pH ~3 and the gel usually contains an organic amine to promote crystallization and structure formation. If lower temperature (100–200°C) is used, crystallization of denser phases including variscite occurs. Note that the synthesis temperature is generally higher and the gel pH is much lower for AlPO materials than for hydrothermal synthesis of aluminosilicate zeolites. Also, the Al/P ratio is typically unity in AlPO materials while the Si/Al ratio is variable in zeolites.

The microporous AlPO materials are usually designated AlPO-*n* where *n* is a number indicating a specific structure type and the *n* designation of new structures is not continuous as for zeolite materials. For example there are no AlPO-1, 2, 3, 4 structures because these preparations turned out to be mixed phases or previously known phases, but there is an AlPO-5 structure. Figure 8 shows typical structures for AlPO-5 with a 12-ring channel and AlPO-11 with a 10-ring channel. AlPO-8 has a somewhat similar structure with a 14-ring structure. In addition to unique structures like these that do not have zeolite analogs, some AlPO structures are analogous to zeolites. AlPO-37 has the faujasite structure, AlPO-34 has the chabazite structure, and AlPO-42 has the A zeolite structure. Some other designations for AlPO materials are also used such as VPI-5 for an 18-ring structure and JDF-20 for a 20-ring structure with hydroxides protruding into the large channel. VPI-5 is the AlPO with the largest empty channel synthesized to date.

The role of an organic cationic or neutral amine in the synthesis gel often seems critical to obtain the desired microporous AlPO structure. However, it is unclear how this actually works. A wide range of organic amines can promote the crystallization of a single material like AlPO-5. Also, a single organic amine can promote crystallization of different AlPO structures depending on the synthesis temperature. After synthesis, the organic agent is generally removed by calcination. The thermal stability of AlPO materials is structure dependent, but is generally as high as 700°C or more.

Another aspect of AlPO materials is that aluminum can have coordination numbers of 4, 5, and 6 rather than being only tetrahedral. This is associated with additional

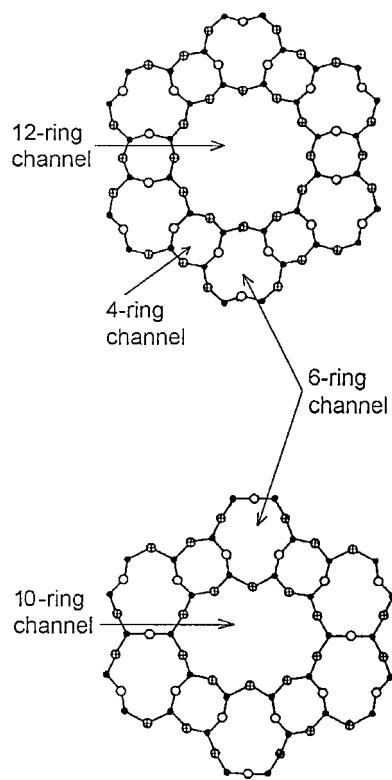


FIGURE 8 Framework structures of AIPO-5 (top) and AIPO-11 (bottom) viewed along the *c* axis where the solid circles are Al or P (T atoms) and the open circles are oxygen atoms; oxygens above the plane of the T atoms in a given ring are represented by \circ and those below the plane by \otimes .

coordination to hydroxide or water in the material. The coordination of aluminum is most directly shown by nuclear magnetic resonance of ^{27}Al since the chemical shifts differ for different coordinations.

A. Silicoaluminophosphates

The ideal AIPO lattice is neutral with no ion exchange capacity. In reality, some small ion exchange capacity occurs due to some hydroxyl group incorporation. However, significant ion exchange capacity may be introduced by adding a silica sol to the synthesis gel. The silicon substitutes more for P than for Al to create a negatively charged framework material termed silicoaluminophosphate (SAPO). Thus, SAPO-5 can be prepared with silicon substituting as much as 20% of the phosphorus. The silicon does not always substitute randomly as isolated sites and may form localized clusters or “islands.”

In general SAPO-*n* structures exist analogous to AIPO-*n* structures. But the added silicate does affect the crystallization process because some SAPO structures

with no AlPO or zeolite equivalent occur. Examples include SAPO-40 (AFR), SAPO-41 (AFO), and SAPO-44 (CHA) where the structure code is given in parenthesis.

The SAPO materials not only have ion exchange capacity but also have acid sites suitable for catalysis. The acid activities are mild and not as high as for zeolites. But this mild activity is optimal for catalysis of certain reactions. The SAPO materials also have high surface areas suitable for sorbent applications. Thus SAPO materials have potential applications in ion exchange, sorption, and catalysis analogous to zeolites.

B. Metal Substituted AIPOs and SAPOs

Several other metal ions (Me) can be added to the synthesis gel and are incorporated into the framework of AIPO and SAPO materials to form MeAPOs and MeAPSOs. The AIPO frameworks seem relatively amenable to such substitutions. This contrasts with zeolites where it is much more difficult to incorporate other ions into the aluminosilicate framework.

Vanadium and titanium can be fairly readily incorporated into AIPOs to form VAPO and TAPO materials by substitution for phosphorus. The amount of framework substitution seems to be as high as several percent. A number of divalent ions including Mg, Co, Fe, Mn, Zn, and Ni among others also appear to participate in framework substitution. It has been argued that these ions primarily substitute for aluminum and for phosphorus. The experimental evidence is ambiguous on which substitution dominates. It is often difficult to tell if framework substitution has actually taken place. The amounts of such substituted metal ion are usually too low to directly determine by powder X-ray diffraction. So, a variety of optical and magnetic resonance spectroscopic methods have been used.

Metal-substituted AIPOs and SAPOs are of interest because of their potential for tailored catalytic activity especially in selective oxidation reactions and for methanol to olefins conversion. This is especially true when the metal ion has variable valence as do most transition metal ions. Thus it is possible that controllable strength redox and acidic activity can be built into a SAPO material. The fact and significance of transition metal ion framework substitution into SAPO materials can be tested by comparison with the same SAPO structure where the same transition metal ion is incorporated by ion exchange, because ion exchange clearly incorporates the ion into a nonframework site. If the properties and spectroscopic signatures are the same for ion-exchanged and -synthesized materials, then synthesis probably has not led to framework incorporation of the metal ion. However, if the spectroscopic signatures are different, there is a reasonable case for framework substitution. Some applications of MeAPSO materials

have been demonstrated but commercial applications have yet to be realized.

SEE ALSO THE FOLLOWING ARTICLES

MATERIALS CHEMISTRY • MESOPOROUS MATERIALS SYNTHESIS AND PROPERTIES

BIBLIOGRAPHY

- Barrer, R. M. (1982). "Hydrothermal Chemistry of Zeolites," Academic Press, London.
- Breck, D. W. (1974). "Zeolite Molecular Sieves," John Wiley, New York.
- Brindley, G. W., and Brown, G., eds. (1980). "Crystal Structures of Clay Minerals and their X-Ray Identification," Mineralogical Society, London.
- Dyer, A. (1988). "An Introduction to Zeolite Molecular Sieves," John Wiley, London.
- Hartmann, M., and Kevan, L. (1999). Transition metal ions in aluminophosphate and silicoaluminophosphate molecular sieves: location, interaction with adsorbates and catalytic properties. *Chem. Rev.* **99**, 635–663.
- Kevan, L. (1992). Applications of electron spin echo modulation to transition metal ions in smectite clays. *Pure Appl. Chem.* **64**, 781–788.
- Meier, W., Olson, D. H., and Baerlocher, C. (1996). "Atlas of Zeolite Structure Types," 4th ed., Elsevier, New York.
- Newman, A. C. D., ed. (1987). "Chemistry of Clays and Clay Minerals," John Wiley, New York.
- Pinnavaia, T. J. (1983). Intercalated clay catalysts. *Science* **220**, 365–371.
- Szostak, R. (1998). "Molecular Sieves," 2nd ed., Blackie, New York.
- Tschernich, R. W. (1992). "Zeolites of the World," Geoscience Press, Phoenix.
- van Bekkum, H., Flanigen, E. M., and Jansen, J.C., eds. (1991). "Introduction to Zeolite Science and Practice," Elsevier, Amsterdam.
- van Olphen, H. (1997). "An Introduction to Clay Colloid Chemistry," 2nd ed., John Wiley, New York.



Nanocomposites and Intercalation Compounds

Christopher O. Oriakhi

Hewlett-Packard Company

Michael M. Lerner

Oregon State University

- I. Intercalation Hosts and Reactions
- II. Preparing Nanocomposites
- III. Nanocomposite Structures
- IV. Properties and Applications

GLOSSARY

Delamination The separation of a layered lattice into thin platelets, in some cases into the smallest units that contain strong intralayer bonds.

Gallery The region between sheets in a layered host occupied by intercalate.

Intercalation A chemical reaction where a molecular or ionic guest species, the intercalate, is introduced into a host lattice, with retention of the strong bonding within the host structure.

Nanocomposite A composite of different compounds where the components are divided into submicron domains in at least one dimension.

Organoclay Clay minerals rendered organophilic, usually by exchange of the intercalated cations for alkylammonium ions.

THE INTRODUCTION of ions or molecules into layered or three-dimensional hosts, a chemical process known as intercalation, has been known since the mid-19th century.

Early researchers in this area included studies clays and graphite as layered hosts, and, later, three-dimensional zeolitic frameworks. In the 1970s, intercalation chemistry was extended to synthetic metal oxides and layered metal chalcogenides. For some time it has been known that molecules can be adsorbed onto layered clay minerals, but since the 1980s interest has grown significantly in obtaining materials that are intercalation compounds with polymeric intercalates. These materials are related to composites where inorganic particles are combined with a polymeric (plastic) phase, and since intercalation results in a nanometer-scale combination in at least one dimension, they have been termed nanocomposite materials. Because interesting or unusual mechanical, electrical, and chemical properties have been demonstrated in many of these new materials, nanocomposites are of great interest in applications such as structural materials, coatings, and electrodes. To date, most synthetic nanocomposites have consisted of binary combinations, those derived from only two components. A progression toward complex materials, such as the biologic nanocomposites comprising shells, bones, or cobwebs, can be expected in the future.

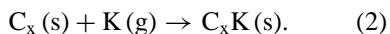
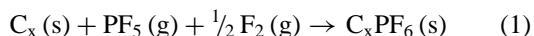
I. INTERCALATION HOSTS AND REACTIONS

While intercalation host structures can be made from sheets, frameworks, chains, tubes, or nanocrystallites, in this article we will mainly describe intercalation of inorganic layer and framework structures and the related work on nanocomposites containing these structures.

A. Layered Hosts

Layered inorganic structures consist of strongly bound sheets held together by relatively weak forces. Since the sheets can be separated without breaking strong bonds, in many cases guest ions or molecules can be intercalated between the layers. The formation of the intercalated gallery can significantly increase the distance between host sheets, and in some cases the complete delamination of the host into isolated single sheets occurs. As described later, the latter phenomenon provides new synthetic and structural possibilities for the preparation of nanocomposites.

The first intercalation reactions of layered hosts to be investigated in detail involved the graphite structure. In 1841, Shafhauti reported that graphite swelled when heated in concentrated H_2SO_4 , in modern terms we understand this to be the intercalation of sulfate and bisulfate anions between the planar carbon sheets that comprise graphite, with concurrent oxidation of these sheets. Graphite is in several ways a prototypical host lattice, its structure of single-layer aromatic hexagonal rings is simple, and yet graphite can intercalate a wide range of guests. Uniquely, graphite allows either anions and cations to intercalate, in the former case the carbon sheets are oxidized and bear a positive charge; in the latter they are reduced and are negatively charged. With graphite, intercalation is therefore a redox reaction, and the reactions require either the use of strong oxidizing or reducing agents, or the application of a potential in an electrochemical cell. The following reactions give examples of graphite oxidation (reaction 1) and reduction (2) to form intercalation compounds. In Reaction 1, the oxidizing agent is fluorine gas, and the intercalated anion is PF_6^- ; in Reaction 2 the reducing agent is potassium metal vapor, which then intercalates as K^+ cations.



Another interesting and special feature of graphite intercalation lies in the phenomenon called staging. Rather than intercalates being evenly distributed throughout the host structure, graphite intercalation compounds usually form ordered arrangements in alternating galleries (Stage 2),

every third gallery (Stage 3), and etc. This ordering of intercalate is not common for other hosts, and it is believed that the high electronic conductivity and flexibility of the carbon sheets are responsible for electronic and mechanical strain energies that favor the formation of the staged arrangements.

The chemistry of the clay minerals was first investigated at about the same time as that of graphite. Interest in the clays developed initially from their technological application as adsorbents, and important phenomenon such as ion exchange and flocculation of particles were investigated in detail. In the 1920s, the first investigations of the structural details for these reactions indicated that clay adsorption chemistry actually involved the intercalation of guest molecules or ions between the aluminosilicate sheet structures of the clay host. Because many clay minerals have lower charged cation substitutions within the sheet structures, for example Al^{3+} for Si^{4+} or Mg^{2+} for Al^{3+} , the aluminosilicate sheets are negatively charged, and cations must be present between the clay sheets to form charge-neutral solid structures. Figure 1 shows the structure of such an example, the clay mineral called montmorillonite. Intercalation reactions with clays are not generally redox reactions, but involve rather the exchange of intercalate cations, or the intercalation or exchange of molecules that interact with either intercalate cations or the clay surfaces.

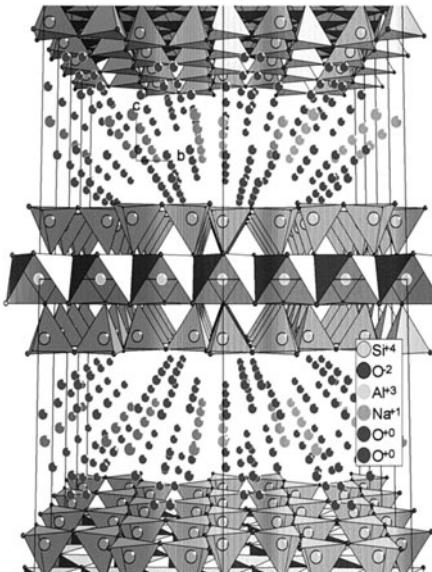


FIGURE 1 Perspective structure of a 2:1 clay mineral such as montmorillonite. The aluminosilicate layers have intercalated Na^+ ions. (Reprinted with permission from Lerf, A. (2000). Intercalation compounds in layered host lattices: Supramolecular chemistry in nanodimensions. In "Handbook of Nanostructured Materials and Nanotechnology" (H. S. Nalwa, ed.), Vol. 5, Academic Press, New York).

In this way, a very wide range of guests can be introduced within the clay galleries. Smectic clays are a class of aluminosilicates that expand in aqueous solutions due to this hydration process; montmorillonite is an example of a smectite clay. Under appropriate conditions, the clay sheets can delaminate entirely to produce colloidal dispersions of single-sheet structures. The reverse reaction, flocculation of the sheets involves reforming a solid, can occur upon addition of an appropriate intercalate, or a change in solution properties such as pH or ionic strength. Similar delamination reactions can also be observed with some nonclay layer structures, some of which are discussed later.

A thermodynamic model for intercalation reactions must consider the relative changes in the lattice and solvation enthalpies for the reacting species. The expansion of the host upon intercalation is generally unfavorable because it results in an increase in the distance for electrostatic interactions between charged host surfaces and intercalate ions; therefore, interlayer expansions are typically seen to be the minimum that is compatible with the steric dimensions of the intercalate. The intercalate is introduced with an orientation and packing arrangement within the galleries that allows for an optimum lattice enthalpy. Neutral polar molecules are also intercalated into many hosts, and, just as in solutions, these molecules act to solvate the intercalated ions.

Under most conditions, clay minerals have hydrophilic, charged surfaces and will not intercalate nonpolar or low-polarity molecules. However, clays can be converted to more hydrophobic materials that are known as organoclays by two methods, by replacing the intercalate cations with long-chain alkylammonium ions or by directly grafting functionalized nonpolar groups onto the clay surfaces. Such organophilic clays are then capable of intercalating nonpolar organic guest molecules.

Layered transition metal dichalcogenides, MX_2 , ($\text{M} = \text{Mo, Nb, Ta, Ti, or Zr}; \text{X} = \text{S, Se}$) form another important class of host materials for intercalation reactions. Intercalation of suitable guests often results in a modification of the composition, electrical, magnetic, optical, and structural properties of the host MX_2 . Guest species include metal ions, organic molecules, organometallic compounds, and polymers. Intercalation of alkali metals, which can be achieved by chemical or electrochemical reactions, is a redox process involving the partial reduction of the host. A typical lithiating agent is n-butyl lithium, as a strong reducing agent it reacts with many layered chalcogenides to form intercalation compounds with Li^+ as the intercalate as indicated in Reaction 3:

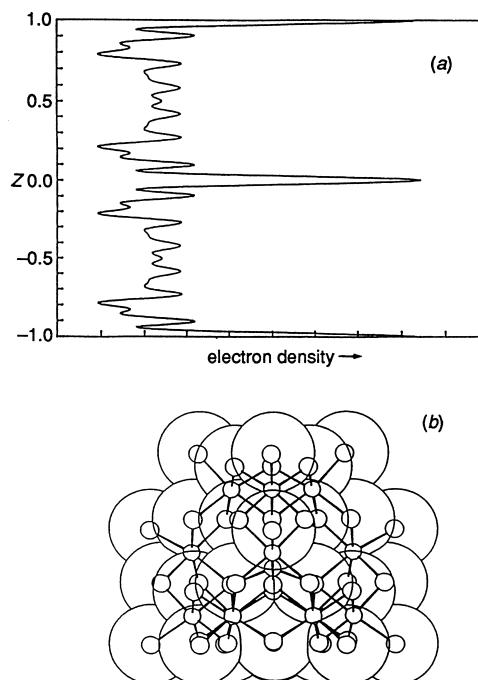
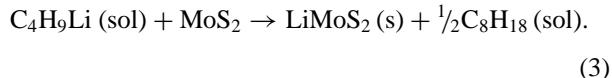


FIGURE 2 (a) One-dimensional electron density projection for the intercalation compound with $[\text{Al}_{13}\text{O}_4(\text{OH})_{24}(\text{H}_2\text{O})_{12}]^{7+}$ guests within negatively charged TaS_2 layers. The large peak at the center and two adjacent peaks indicate the position of the Ta and S layers in the host, respectively. In (b), the structure of the complex aluminum cation intercalate is shown. (Reprinted from Nazar, L. F., and Jacobson, A. J. (1994). *J. Mater. Chem.* **4**, 1419–1425. Reproduced by permission of the Royal Society of Chemistry.)

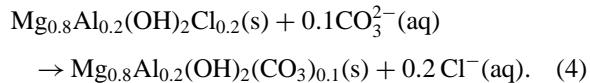
More complex intercalate guests can also be introduced. Figure 2 shows the one-dimensional electron density map for a compound prepared by the intercalation of $[\text{Fe}_6\text{S}_8\text{P}(\text{C}_2\text{H}_5)_3]^{2+}$ cation clusters between TaS_2 layers. The large cations are inserted by first delaminating the host structure and reacting the colloidal dispersion with a solution containing the cationic species. As with the clay minerals, intercalation compounds can be prepared by direct insertion, ion-exchange reactions, or a delamination and flocculation method. In these syntheses, solvent may be co-intercalated along with the cation. Table I indicates some common layered host structures for intercalation chemistry. Other important classes of intercalation hosts include metal phosphorus trichalcogenides, MPX_3 , ($\text{M} = \text{Mg, Mn, Fe, Co, Cd, Ni, V, and Zn}; \text{X} = \text{S, Se}$); transition metal oxyhalides, MOX ($\text{M} = \text{Cr, Fe, Ti, and V}; \text{X} = \text{Cl, Br, and I}$); transition metal oxides such as MoO_3 , V_2O_5 , and MO_4 , ($\text{M} = \text{Mo, Ta, Nb, and V}; \text{X} = \text{As, P, S}$); and metal phosphates and phosphonates.

Layered double hydroxides (LDHs), also known as anionic clays, form a unique class of host materials for intercalation. These clays are anion-exchangers and occur naturally: many new synthetic examples have

TABLE I Layered Inorganic Host Structures for Intercalation Chemistry

Layered host lattice	Representative examples
Elemental	Graphite, black phosphorous
Chalcogenides	MX_2 ($\text{M} = \text{Ta}, \text{Ti}, \text{Zr}, \text{Nb}, \text{Mo}, \text{W}; \text{X} = \text{S}, \text{Se}, \text{Te}$)
Phosphorous trichalcogenides	MPX_3 ($\text{M} = \text{Cd}, \text{Fe}, \text{Mn}, \text{Ni}, \text{Zn}; \text{X} = \text{S}, \text{Se}$)
Oxides	LiMO_2 ($\text{M} = \text{Co}, \text{Ni}$), MoO_3 , $\text{Mo}_{18}\text{O}_{52}$, V_2O_5
Phosphates	$\text{M}(\text{HPO}_4)_2$ ($\text{M} = \text{Ti}, \text{Zr}, \text{Ce}, \text{Sn}$); $\text{VOPO}_4 \cdot 2\text{H}_2\text{O}$
Halides	RuCl_3 , PbI_2 , ZnNCl
Oxy-halides	MOX ($\text{M} = \text{Ti}, \text{V}, \text{Al}, \text{La}, \text{Cr}, \text{Fe}; \text{X} = \text{Cl}, \text{Br}$)
Silicides	CaSi_2
Titanates/niobates	KNbO_3 , $\text{K}_4\text{Nb}_6\text{O}_{17}$, $\text{K}_2\text{Ti}_4\text{O}_9$, $\text{H}_2\text{Ti}_3\text{O}_7$, KTiNbO_5
Layered double hydroxides	$[\text{Mg}_6\text{Al}_2(\text{OH})_6]\text{CO}_3 \cdot 4\text{H}_2\text{O}$ (hydrotalcite) $[\text{M}_{1-x}\text{M}'_x(\text{OH})_2]\text{An}_{x/n} \cdot y\text{H}_2\text{O}$ $(\text{M}' = \text{Ca}, \text{Mg}, \text{Co}, \text{Fe}, \text{Ni}, \text{Mn}, \text{Zn}, \text{Li}; \text{M}' = \text{Al}, \text{Cr}, \text{Fe}, \text{V}; \text{An} = \text{anion with charge n-})$
Smectite clays	$\text{Ca}_{0.35}[\text{Mg}_{0.7}\text{Al}_{3.3}](\text{Si}_8)\text{O}_{20}(\text{OH})_4$ (montmorillonite) $\text{Na}_{0.6}[\text{Li}_{0.6}\text{Mg}_{5.4}](\text{Si}_{8.0})\text{O}_{20}(\text{OH}, \text{F})_4$ (Hectorite) $\text{Al}_4\text{Si}_4\text{O}_{10}(\text{OH})_8$ (kaolinite)
Silicic acids	$\text{H}_2\text{Si}_2\text{O}_5$; $\text{H}_2\text{Si}_{14}\text{O}_{29}$

also been prepared. The LDH structures contain positively charged layers based on the charge neutral brucite ($\text{Mg}(\text{OH})_2$) structure with higher valent substitutions, such as Al^{3+} for Mg^{2+} . In the LDH general formula $[\text{M}_{1-x}^{2+}\text{M}_x^{3+}(\text{OH})_2]^{x+}[\text{A}^{n-}]_{x/n}(\text{H}_2\text{O})_y$, where $\text{M}^{2+} = \text{Ca}, \text{Mg}, \text{Ni}, \text{Co}, \text{Cu}, \text{or Zn}; \text{M}^{3+} = \text{Al}, \text{Cr}, \text{or Fe}$. A indicates the charge-compensating anion that resides between the host layers along with water. Intercalation compounds and nanocomposites based on LDH structures can be prepared by anion-exchange reactions starting with an LDH, by templated or direct synthesis where the desired anion is incorporated into the LDH growth process, or by rehydration of calcined LDH in the presence of the desired organic anion. Reaction 4 shows a typical anion exchange process for an LDH:



The carbonate anion is very stable and readily displaces simple monovalent anions when the compounds are exposed to the CO_2 in air, therefore the chloride intercalate, and most other singly charged anion intercalate phases, can only be prepared in a rigorously air-free atmosphere.

B. Three-Dimensional Inorganic Structures

Most three-dimensional intercalation hosts are inorganic oxides that contain channels with diameters of about 4–

6 Å. These hosts are commonly called zeolites and were first prepared in the 1950s. About two or three dozen naturally occurring zeolite minerals are known, many more have been prepared synthetically. As with aluminosilicate clays, zeolites have negatively charged frameworks, and therefore contain cations intercalated within the internal pores and cavities to form charge neutral solids. Zeolites undergo ion exchange reactions or the intercalation of solvent molecules. The ordered channels and cavities result in large internal surface areas and allow for the rapid intercalation of small ions or molecules from either a gas or liquid phase. Since these inorganic frameworks have strong bonds in three dimensions, they can expand only slightly, therefore, the pore diameters limit the maximum intercalate diameter. This results in far more selective adsorption properties than is observed with clay hosts. For example, zeolite 4A can absorb water and some other small molecules, but larger molecules, such as branched hydrocarbons, cannot fit through the octagonal windows that form the smallest diameter of the connecting pore structures in this zeolite.

Zeolites are usually prepared by the nucleation of aluminate and silicate-containing solutions around organocations, so the incorporation and exchange chemistry of such cations affects the growth kinetics and structure of synthetic zeolites. For this reason, the structure and chemistry of organocation intercalates have been studied in detail. Complex organic or organometallic species, such as metallophthalocyanines and metallocenes, have also been intercalated into zeolites either via adsorption, ion exchange, or template growth. Although the large majority of zeolites are based on an aluminosilicate framework, some have been prepared with other oxide frameworks (one important example being AlPO_4).

II. PREPARING NANOCOMPOSITES

The thermodynamic driving forces that result in intercalation chemistry are generally similar for nanocomposites, where the guests are polymers. Charged polymers, or ionomers, can be incorporated and form electrostatic bonds to oppositely charged inorganic surfaces. Polar or functional polymers, as with smaller molecules, might solvate intercalate ions, although recent studies have suggested that the interactions with the charged host surfaces may be as, or more, significant. Nonpolar polymers require the modification of the host to obtain a hydrophobic (organophilic) host surface or intercalate gallery. Polymer adsorption and intercalation can also be driven by the entropically favorable displacement of smaller molecules. **Table II** provides a glossary of the polymer structures discussed in this article.

TABLE II Glossary of Polymers in This Article

Name	Abbrev.	Structural repeat unit
Polyaniline	PANI	
Polypyrrole	PPY	
Polythiophene	PTH	
Poly(p-phenylene)	PPP	
Poly(p-phenylene vinylene)	PPV	
Poly(phenylene sulphide)	PPS	
Poly(p-phenoxyphenylene sulfide)	PPPS	
Polystyrene	PS	
Poly(styrene sulfonate)	PSS	
Poly(vinyl sulfonate)	PVS	
Poly(vinylacetate)	PVA	
Poly(acrylonitrile)	PAN	
Poly(acrylamide)	PAA	
Poly(ethylene oxide)	PEO	
Poly(vinyl alcohol)	PVOH	
Poly(ethylenimine)	PEI	
Poly(methyl methacrylate)	PMMA	

continues

TABLE II (continued)

Polyacetylene	PA	
Polyethylene terephthalate	PET	
Poly(ether ether ketone)	PEEK	
Poly(dimethylsiloxane)	PDMS	
Poly(caprolactone)		
Poly(caprolactam) (Nylon 6)		
Poly(vinyl carbazole)		

Although the energetic considerations are often similar for the intercalation of small and polymeric intercalates, the kinetic restrictions that arise due to the limitations of polymer diffusion usually require very different synthetic strategies to be employed in the preparation of nanocomposites. Topotactic or direct reactions, that is, intercalation reactions where the host structure is preserved throughout the reaction and guests enter via pores or intersheet galleries, are well suited to small intercalate ions or molecules but are often not practical for polymers. On the other hand, the conventional method of preparing composites by mechanical mixing of the two constituent phases is not often effective in producing nanoscaled combinations for inorganic hosts. Therefore, significant synthetic challenges have arisen in preparing nanocomposites, and in many cases new methods have been developed for these materials. In order to avoid the need for solid-state diffusion of large, polymeric structural components, most syntheses have involved either (1) the intercalation and subsequent *in situ* polymerization of the guest within the inorganic host framework, or (2) the growth or dispersion of the inorganic component in polymer-containing solutions. In both cases, the nanocomposites are obtained by growing one component about (or within) the other. The greatest number of studies, and synthetic strategies, have been developed for zeolitic and layered nanocomposites, and these will therefore be described in the most detail. Table III presents a short list of some nanocomposites

TABLE III Selected Nanocomposites with Two- and Three-Dimensional Inorganic Components

Polymer	Layered hosts	Preparation method
PANI	MoO ₃ , FeOCl, RuCl ₃ , V ₂ O ₅	<i>In situ</i> polymerization
PPY	V ₂ O ₅ , FeOCl, VOPO ₄	<i>In situ</i> polymerization
PTH	V ₂ O ₅	<i>In situ</i> polymerization
PPV	MoO ₃	Thermal conversion
PPV	Clay	Chemical conversion
PANI	MoS ₂	Exfoliation/adsorption
PEO	MoS ₂ , MoSe ₂ , TiS ₂ , MnPS ₃ , CdPS ₃ NbSe ₂ , MoO ₃ , clay, graphite oxide	Exfoliation/adsorption
PEI	MoS ₂ , MoSe ₂ , TiS ₂ , clay	Exfoliation/adsorption
Poly(phosphazene)	Clay	Exfoliation/adsorption
Nylon-6	Clay Organoclay	<i>In situ</i> polymerization <i>Melt</i> intercalation
PSS	Ca-Al-LDH, Mg-Al-LDH, Zn-Al-LDH	Template synthesis
Poly(acrylic acid)	Ca-Al-LDH, Mg-Al-LDH, Zn-Al-LDH	Template synthesis
PVS	Ca-Al-LDH, Mg-Al-LDH, Zn-Al-LDH	Template synthesis
Polymer	Framework (3-D) hosts	Preparation method
PAN	Zeolite-Y (NaY), mordenite, MCM-41	<i>In situ</i> polymerization
Poly(furfuryl alcohol)	Zeolite-Y	<i>In situ</i> polymerization
PMMA	Zeolite-Y, mordenite, ZSM-5, MCM-41, I	<i>In situ</i> polymerization
PS	MCM-41	<i>In situ</i> polymerization
PVA	MCM-41	<i>In situ</i> polymerization
Poly(ferrocenyl silane)	MCM-41	<i>In situ</i> polymerization
PANI	Mordenite, zeolite-Y, MCM-41	<i>In situ</i> polymerization
PPY	Faujasite, mordenite, MCM-41, zeolite-L	<i>In situ</i> polymerization
PMA	Mordenite, SAPO-5, ZSM-5	<i>In situ</i> polymerization

prepared using two- and three-dimensional inorganic components.

A. *In Situ* Polymerization

A common method for incorporation of polymers within layered or three-dimensional hosts involves the intercalation of a monomeric precursor followed by the *in situ* polymerization within the host itself. This method avoids the need for diffusion of macromolecules, and can be especially useful if the target polymer is not soluble. Several important examples of this method have been reported for polymers with conjugated backbones, where the monomers can be dissolved or reacted in the gas phase, but the polymers are either completely insoluble, or must be dissolved into solvents that are incompatible with either the intercalation reaction conditions, the stability of the host, or both.

In situ polymerization can be accomplished by several methods. If the host is sufficiently transparent, ultraviolet radiation can be used to effect polymerization reactions where appropriate. Other methods involve the use of a

polymerization agent, for example, polyaniline (PANI) nanocomposites with mordenite and zeolite Y have been prepared by the intercalation of aniline and subsequent oxidative polymerization using a persulfate-containing solution.

Alternately, the polymerization reagent can be another intercalate, as where pyrrole has been introduced into a Cu(II)-exchanged fluorohectorite. Polypyrrole (PPY) and polythiophene (PTH) have also been formed within zeolite pores by the ion exchange of intercalate cations with Cu(II) or Fe(III) prior to monomer intercalation and polymerization. Finally, the host can itself be the source of the polymerization agent, an example is where pyrrole is polymerized by Fe(III) in the layered FeOCl host. In another example, spectroscopic studies on methylacetylene that was polymerized within protonated zeolites suggest that some monomeric precursors may polymerize spontaneously due to catalysis by acid sites present on the host surface.

Layered structures can expand along the stacking direction to accommodate polymeric intercalates, but three-dimensional frameworks are fairly rigid and therefore have

restricted pore dimensions that limit monomer access and the types of polymers that can be accommodated. The linear backbones of many conjugated polymers are therefore especially well suited for such hosts.

An issue associated with the *in situ* polymerization method is how to control and determine the polymer properties when the polymerization occurs under very different conditions from those used in more traditional and well-understood polymerization reactions. In some cases, the polymer can be extracted from the nanocomposite and then characterized by conventional methods in solution. The extraction of PANI after *in situ* polymerization within V₂O₅ showed a product molecular weight of 14,000, about half that obtained for the *ex situ* polymerization using similar conditions. Extraction of PANI from layered polyphosphates gave an M_w of 3000–9000, and these shorter polymer chains obtained were also found to be partially crosslinked.

Another method of evaluating the polymer has been to etch away the inorganic oxide framework with dilute HF, which under controlled conditions can dissolve metal oxides without reacting with organic polymers. Following this step, the polymer residue can be redissolved and evaluated by conventional methods. By this method, the *in situ* free-radical polymerization of acrylonitrile within mordenite and NaY zeolites was determined to produce polymeric polyacrylonitrile chains with molecular weights as high as 19,000. Figure 3 provides a schematic view of this reaction.

Some zeolite-like mesoporous structures, with pore sizes between 15 and 100 Å, have also been used to prepare nanocomposites by *in situ* polymerization. An example is the hexagonal, mesoporous MCM-41 structure first characterized by Mobil researchers in 1992. Using

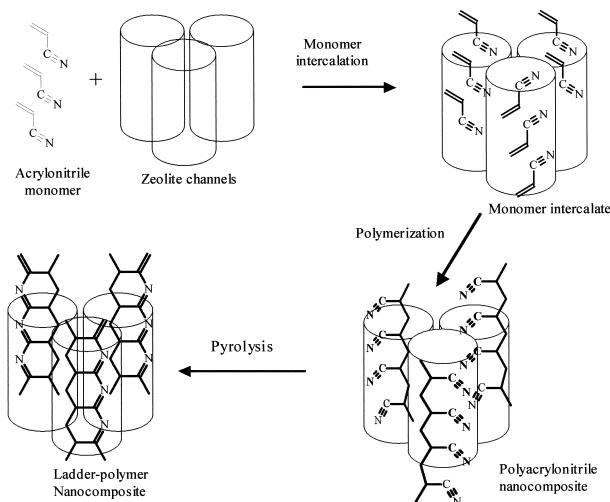


FIGURE 3 Polymerization of acrylonitrile within the channels of a zeolite host.

this host, a number of organic monomers, including acrylonitrile, aniline, methyl methacrylate, styrene, and vinyl acetate, have been intercalated and polymerized to afford the corresponding nanocomposite materials.

Low-polarity monomers can also be intercalated into inorganic hosts and polymerized. The host in this case must be modified with an organocation (as described earlier) in order to allow the nonpolar monomer to intercalate. Many structural polymers have low polarities, and are therefore of great interest as nanocomposite components. An important example lies in the nanocomposites formed with nylon-6 and alkylammonium-modified montmorillonite. Additional examples of polymer incorporated using this method include poly(ϵ -caprolactone), polyimide, epoxy resins, and polystyrene (PS).

B. Direct Intercalation

In a few cases, it has proven possible to intercalate polymers by the same topochemical reactions used for small molecules. Nanocomposites with poly(ethylene oxide) (PEO) and montmorillonite, or PS and poly(dimethylsiloxane) with organoclays, have been formed simply by heating mixtures containing the polymers and inorganic hosts above the polymer melting or glass-transition temperature. PEO has also been shown to intercalate directly by heating the polymer in pelletized mixtures with layered silicates, MPS₃ and MoO₃ host structures.

This direct method has the significant advantage that solvents are not required, and therefore standard processing methods such as extrusion can be used to obtain the nanocomposite materials in useful forms. While the reaction mechanism is not fully understood, PEO is an elastomeric polymer that has facile segmental motion at the reaction temperatures used. It may be that this allows for a sufficient rate of polymer diffusion. A study on polystyrene intercalation kinetics found that the mobility of polymer chains within the nanocomposites is comparable to that in the molten polymer. Some studies indicate that host delamination, or the separation of the layered hosts into very thin or single-sheet particles, occurs along with polymer incorporation, which suggests that the exfoliation of the layered hosts provides an alternate mechanism of nanocomposite formation in these solvent-free conditions.

Several companies, including Allied Signal, Ford Motors, ICI, and Toyota have developed or are developing nanocomposite technology based on clays combined with polymers such as nylons, polypropylene, polyester, or epoxies. A great majority of these nanocomposites are based on modified clay mineral host materials (organoclays). The mass fraction of clay in the nanocomposites can vary from 2 to 70%. When the clay mass fraction is lower than 20%,

an exfoliated amorphous material, in which single silicate sheets are dispersed within the polymer matrix is usually obtained. When the mass fraction of clay is greater than 20%, an intercalated nanocomposite containing ordered multilayer of polymer and silicate sheets separated by a few nanometers is obtained. It is interesting to note that, in general, only a small inorganic mass fraction, in the range of 2–10%, is required in order to obtain significant changes in the physical and chemical properties of the nanocomposites.

C. Exfoliation–Adsorption

Figure 4 provides a schematic representation of the exfoliation–adsorption method. In this process, the layered host sheets are delaminated and dispersed into colloidal suspensions containing single-sheet host platelets. This step avoids the diffusion limitations of the topotactic method of polymer intercalation, because the direct interaction of all sheet surfaces with species in the solution is possible. When the soluble polymer interacts with a clay platelet, it may adsorb onto the individual sheets, and a nanocomposite structure is then obtained by the subsequent aggregation of such assemblies. The aggregation step may be spontaneous for some intercalates, or induced by changes in pH, solvent polarity, or concentration. This method has the advantage of using preformed, and often well-characterized, polymers rather than forming the polymer under conditions that can be difficult to control.

The first reports on the intercalation of macromolecular species were studies of polymer adsorption by colloidal

clay minerals. Many water-soluble polymers, including both polar and cationic polymers, adsorb onto clay surfaces and remain attached to the surfaces to form intercalate galleries between the clay sheets when the solid aggregates. In the 1980s, metal disulfides were also found to form single-sheet colloidal suspensions when their lithiated forms were rapidly hydrolyzed.

As with the clays, these layered sheet colloidal dispersions can react with organic species to form ordered intercalation compounds or nanocomposite structures. Another layered structure, MoO_3 , has also been shown to form nanocomposites by this method. It seems this approach can be even more general in the future. A layered sheet surface to charge ratio of approximately $40\text{--}120 \text{ \AA}^2/\text{e}^-$ has been proposed as appropriate for stabilizing colloidal dispersions for inorganic hosts, and thus for forming reactive precursors that can lead to nanocomposites.

When polymers cannot dissolve in solvents that are compatible with the colloidal dispersion containing the inorganic sheets, similar methods can be employed if the polymer can be obtained in a finely dispersed form. Nanocomposites containing PPY, PTH, and poly(p-phenoxyphenylene sulfide) within montmorillonite galleries were formed by reaction of a latex (a solution containing suspended polymer nanoparticles) with a solution containing the delaminated host.

D. Thermal or Chemical Conversion

In many cases, derivatives of otherwise intractable polymers have been developed that can be more readily processed, for example, into thin films, and then converted into the desired polymer by a chemical reaction or thermolysis. Some important examples include the conjugated polymers poly(p-phenylene), polyacetylene, poly(phenylene sulfide), and poly(phenylene vinylene) (PPV) or thermoplastics such as poly(ether ether ketone) and poly(arylene sulfide ketone). Each of these polymers has been obtained in medium or high molecular weight by the conversion of a soluble polymeric precursor. This approach has also been adapted to form nanocomposite structures containing these polymers. The precursor polymers are solubilized, and then reacted, according to the exfoliation–adsorption method, and the product is subsequently converted by chemical or thermal means into one containing the target polymer.

An example of the use of this method lies in the preparation of a PPV/ MoO_3 nanocomposite by interaction of colloidal MoO_3 sheets with a solution containing poly(xylylene dimethylsulfonium). This cationic precursor polymer readily aggregates with the anionic sheets to form an ordered nanocomposite structure. Subsequently, the controlled *in situ* thermolysis of the polymer at 250°C

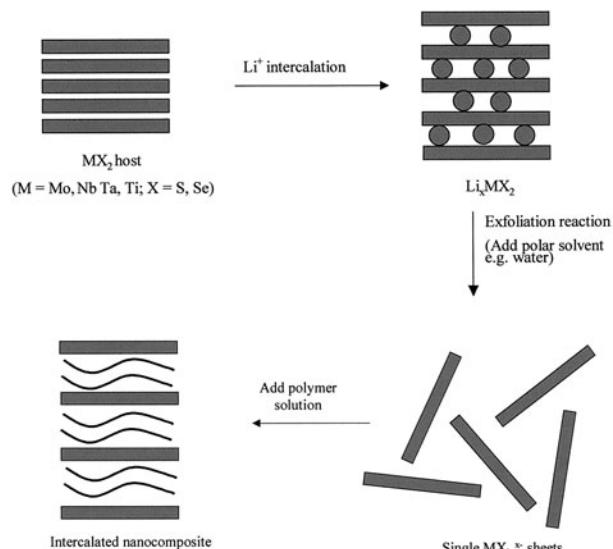
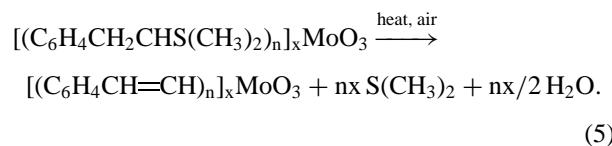
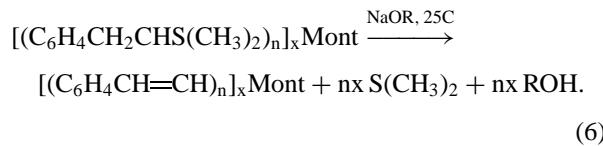


FIGURE 4 Preparation of layered nanocomposites using the exfoliation–adsorption method.

eliminates dimethylsulfide to form the desired product as shown in Reaction 5:



In an alternate method to prepare a nanocomposite with the same polymeric component, the chemical conversion of the same precursor in montmorillonite (Mont) can be accomplished by reaction with base as shown in Reaction 6:



The method shown in Reaction 6 may lead to higher-molecular-weight polymer in the nanocomposite, because the temperature required for the conversion step is close to the limit of thermal stability for PPV.

E. Sol-Gel or Template Reactions

Nanocomposites with either ordered or amorphous structures can be obtained from solutions containing both a polymer and a metal alkoxide precursor to the inorganic component. Sol-gel chemistry is used to prepare solid oxides by the controlled hydrolysis of alkoxide derivatives of one or more inorganic precursors to form a sol containing oligomeric units of metal oxide precursors. Subsequent steps involve gel formation, densification, drying and annealing to obtain the final product. When a soluble polymer or monomer is included in the sol, and the processing steps limited to low temperatures that will not degrade the polymeric component, nanocomposites can be prepared by this sol-gel method. Such reactions can produce very homogeneous materials, and also permit some control over the product particle size and morphology (shape). While in the gel form, the material can often be cast or spin-coated to form thin films. Inorganic glasses with polymeric inclusions form an important area of research: a typical example is the incorporation of methylmethacrylate within silica by sol-gel processing and subsequent polymerization.

Ordered structures containing layered double hydroxide (LDH) sheets and anionic polymers have been obtained by a related method involving the precipitation of metal salts from a basic solution. As an example, a nanocomposite with 10 Å thick layers of poly(vinyl alcohol) between Ca-Al-LDH sheets was formed by increasing the pH of a polymer solution containing the sheet-forming

metal cations in appropriate stoichiometry. Using the same method, poly(styrenesulfonate), poly(vinylsulfonate), and poly(vinylacrylate) have also been shown to form polyanion layers in LDH hosts.

III. NANOCOMPOSITE STRUCTURES

A. Structures with Three-Dimensional Inorganics

In Fig. 3, the structure of a polymer with a linear backbone within zeolite channels is represented schematically. In such materials, if all the polymer is contained within the zeolitic internal volume, the electronically conducting chains are isolated from one another, and low bulk electrical conductivities are observed. In contrast to the zeolites, mesoporous hosts may accommodate more than one polymer chain per channel. MCM-41 is a mesoporous aluminosilicate that contains 3-nm-diameter channels. The incorporation of PANI by *in situ* polymerization produces high-polymer molecular weights (up to 35,000), and compositional, modeling and sorption data indicate the presence of several polymer chains per channel. Contactless conductance measurements also have been used to establish that the polymer chains are in a conducting form.

Conducting polymers have also been intercalated into the nanopores and mesopores of frameworks comprising alumina or polycarbonates, again by *in situ* polymerization using chemical oxidants or electrooxidation. One interesting result has been the formation of tubular polymer structures due to nucleated growth of polymer chains at the walls of the mesopores. At long reaction times, some conjugated polymers form tubules that close to form solid fibrils, while others do not. Further processing of these tubular structures has shown that they may be used to immobilize enzymes, with a potential application in the controlled delivery of drugs.

B. Structures with Layered Inorganic Hosts

When nanocomposites are prepared with a predominant inorganic content, layered structures are often obtained where one or two molecular layers of polymer, with thicknesses of 3–10 Å, are intercalated between inorganic host sheets. These structures can be well ordered along the stacking direction, and the gallery expansion then determined by X-ray diffraction methods. If, on the other hand, the polymer component predominates, exfoliated nanocomposites are often obtained where single inorganic layers are dispersed in a continuous polymer matrix. These exfoliated nanocomposites may show regions

of polymer crystallinity depending on the bulk polymer. The dispersed inorganic sheets may show no significant ordering, in which case the interlayer separation is related to a statistical average concentration of host layers, which could mean distances of as much as hundreds of nanometers. In some cases, even highly dispersed inorganic sheets can show long-range correlations in orientation, especially when the ordering is induced by processing conditions such as pressure or rate of product film deposition. In Fig. 5, a transmission electron micrograph shows the structure of a nanocomposite formed from montmorillonite and a polypropylene–maleic anhydride copolymer. In the figure, exfoliated single, double, and triple layers of host, as well as a multilayer tactoid particle, can all be observed within the continuous polymer phase. While the orientation and organization of polymers is restricted by the narrow pore dimensions of zeolitic hosts, the expanded galleries in layered hosts means that many polymer chain orientations are possible. The structural characterization by diffraction or microscopy most often indicates only the amount of expansion along the stacking direction, and in these cases, the limited size of ordered domains can give relatively high uncertainties for even

these data. Structural characterization can also include methods such as electron microscopy, solid-state NMR, elemental and thermal analyses, and other spectroscopic methods.

Although the diffraction data may be limited to indicate the expansion between host layers after intercalation of the polymer, this limited information can often provide sufficient information to suggest a structural model with a specific polymer orientation within nanocomposite galleries. When PANI is obtained in a highly charged, cationic form and intercalated within clays, V_2O_5 , or MoO_3 , the polymer forms layers 5–6 Å thick between host sheets. This dimension matches that calculated for a single polymer chain with the N or NH groups oriented towards the host surfaces. When the polymer is intercalated with a lower positive charge, the resulting nanocomposites often have a polymer layer of only 3.5–4.5 Å thickness. This is certainly too small to permit the same orientation as above, but does correspond well with the expected dimension if the polymer's conjugated planes are oriented parallel to the inorganic surface.

The first arrangement may allow more effective interactions between the cationic sites on the polymer and the negatively charged inorganic sheets, and be favored for a higher concentration of cationic sites present along the polymer chain. With such limited data, however, unambiguous models cannot always be assigned. The larger expansion, for example, could also correspond to a model where polymer chains are twisted along the chain axis, or even one where the polymer chains are completely disordered (although the chains must be restricted to lie within the same approximate layer in order to fit within the observed dimensions). For almost all conjugated polymers, the small gallery expansions obtained have indicated that only monolayers are intercalated between host sheets. Since these materials are produced by a variety of synthetic methods, this result seems to indicate a thermodynamically favorable state rather than a kinetic reaction limit for polymer uptake.

Other polymers are known to form bilayers within the intercalate galleries. Poly(ethylenimine) forms 4.0–4.5 Å layers in several inorganic hosts, as does PEO if it is combined with a variety of layered hosts with the appropriate, polymer-limited, stoichiometries. Steric considerations indicate again that in both cases only polymer monolayers can exist between inorganic sheets. However, when PEO is combined with many layered hosts in higher ratios, the interlayer expansion increases to 8–9 Å. One early model for the larger gallery dimensions suggested the presence of helical PEO chains similar to those observed in crystalline PEO, but there is now spectroscopic and diffraction evidence to confirm that these galleries have a bilayer polymer structure. No ordered structures with

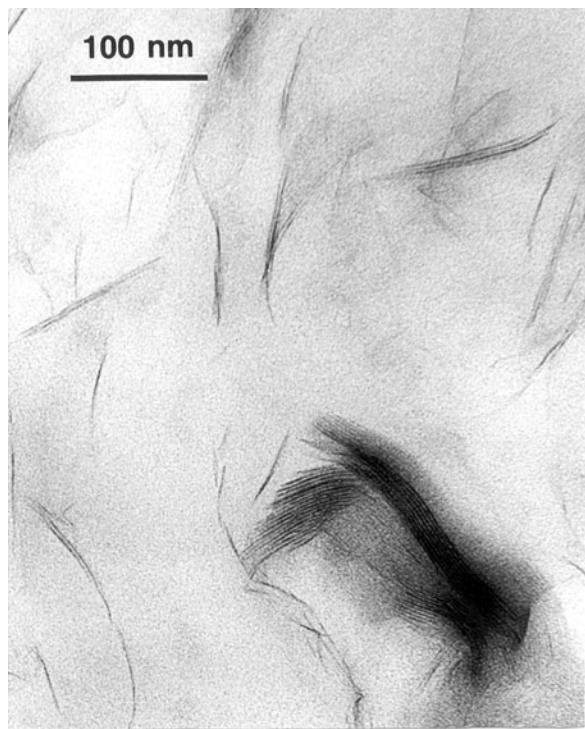


FIGURE 5 TEM of a nanocomposite formed from montmorillonite and a polypropylene–maleic anhydride copolymer. (Reprinted with permission from Gilman, J. W., Jackson, C. L., Morgan, A. B., Harris Jr., R., Manias, E., Giannelis, E. P., and Wuthenow, M. (2000). *Chem. Mater.* **12**, 1866–1873. Copyright 2000 American Chemical Society.)

expansions consistent with trilayers or larger ordered layer structures for these polymers have been confirmed to date.

C. Other Inorganic Structures

There are few examples of one-dimensional inorganic structures that form ordered nanocomposites, primarily because a lower dimensionality does not make the retention of the original host organization as likely. There are, however, several chain structures, such as AFeS₂ (A = alkali metal), containing appropriate charge densities to form colloidal dispersions and therefore amenable to nanocomposite syntheses. Other potential inorganic chains include 2HgS.SnBr₂ (where the SnBr₂ dimer is inserted between the helical chains) and molybdenum and vanadium phosphates with linear chain structures.

The Chevrel-phase ternary molybdenum chalcogenides of general formula AMo₃X₃ (where A = alkali metal, In, or Tl, and X = S or Se) form host-guest complexes, and the formation of a polymer complex between LiMo₃Se₃ and PEO has been reported. This material appears to contain isolated metallic inorganic chains within a polymer matrix.

Semiconductor and metal nanocrystals may be either coated with polymers or included within a continuous polymer phase in order to obtain nanocomposites. Conducting polypyrrole coatings have been reported on silica, SnO₂, and other metal oxides. Several groups have reported dispersions of nanocrystalline metal sulfides within polymer matrices; examples include PbS/PEO, CdS/polyvinylcarbazole, and CdSe or CuS/polyacrylonitrile. The nanocomposites in these cases are prepared by coprecipitation reactions.

IV. PROPERTIES AND APPLICATIONS

Because of the atomic-level combinations of different structure types, some nanocomposites may show properties predominated by the interfacial interactions, and others may exhibit the quantum effects associated with nanodimensional structures. Nanocomposites can therefore exhibit properties that are distinct from the pure component phases, and, perhaps more notably, properties may be obtained that are distinct from the isocompositional macro- or microcomposite materials. Additionally, many nanocomposites are potentially useful because of the combination of desired properties that are obtained from different components. In many cases where new materials are examined for real-world applications, it is often not a single novel property that is of interest, but rather a set of properties that was not readily obtained by other combinations of materials. Studies on nanocomposites have

included fundamental studies of electronic, mechanical, thermal, chemical and optical properties, along with related research and development for real applications in some of these areas.

A. Electronic/Electrochemical Properties

The electrical conductivities of nanocomposites that are derived from inorganic hosts and conjugated polymers can increase by up to several orders of magnitude relative to those obtained for intercalation compounds of the corresponding monomers. This provides evidence for the formation of conducting polymer chains in the nanocomposite. The electronic conductivities are typically higher for layered hosts than for the zeolite-based materials, although the values obtained in both cases are usually lower than those of the pure polymers (compare for example: $\sigma < 10^{-8}$ Scm⁻¹ for PANI/zeolite, 10^{-3} for PANI/MoO₃, and 10^{-2} – 10 for bulk PANI powder). The properties of conjugated polymers, and especially PANI, are closely dependent on preparative conditions. The *in situ* polymerization reaction within the host galleries provides a ready and effective means of nanocomposite formation, but can also lead to low or disperse molecular weights. However, due to the air sensitivity and poor thermal and mechanical properties of many conducting polymers, the conducting nanocomposites are of interest in applications such as anti-static coatings where the polymers alone may show insufficient processability or environmental stability.

At the present time, most primary and all secondary Li or Li-ion batteries use a liquid organic electrolyte solution. This greatly contributes to the fire and explosion hazards of such cells, and also limits to some extent the energy storage in these devices. Due to the established high ionic conductivities associated with PEO-salt complexes, there has been wide interest in the application of polymeric electrolytes in such cells. The application of nanocomposites in this area may prove extremely useful in controlling the mechanical, thermal, and chemical stabilities of such electrolytes. An added potential advantage lies in the fact that such nanocomposites often show higher transport numbers for Li⁺, or are true single ion (Li⁺) conductors. In Li-cell liquid electrolytes, anion transport is not desirable as it leads to electrode polarization during cell operation. Since the negative charge in many nanocomposites is located on the immobile, extended inorganic component, Li⁺ accounts for all or a large fraction of total ionic conductivity. A number of reports on ionic conductivity in PEO/montmorillonite nanocomposites show that they may prove useful for solid electrolytes. Some related oligo- or polyethers, including a random copolymer of PEO and poly(methylene oxide),

and poly(bis-methoxyethoxyethoxy phosphazene), have also been explored as components for nanocomposites in this application.

B. Thermal Stability and Flame Retardancy

Materials with high thermal stabilities are required for numerous applications. In general, organic polymers will thermally degrade at relatively low temperatures as compared with inorganic materials. However, the thermal stability of organic polymers can be significantly improved by forming nanocomposites with inorganic solids. This observation was first made by Blumstein in 1965 when he studied the thermal behavior of nanocomposites derived from clay and poly(methylmethacrylate) (PMMA). The increased thermal stability was ascribed to the confinement and restricted thermal motion of the PMMA within the clay galleries. Recent studies on the thermal behavior of nanocomposites derived from some structural polymers and layered double hydroxides have confirmed the improved thermal stability compared to the parent phases.

Flammability is another important issue for many applications, and one of particular concern for some important classes of polymers. Again, by forming nanocomposites with nonflammable inorganic structures, the flammability of the resulting material can be significantly reduced. Studies have demonstrated that nanocomposites prepared from nylon family, epoxy, polystyrene, or vinyl ester polymers and clays exhibit reduced flammability compared to pure or conventionally filled polymers. The use of Cone calorimetry, which simulates fire-like conditions, indicated that both the heat release rate (HRR) and carbon monoxide output were reduced when the polymers were incorporated into nanocomposite structures. **Figure 6** compares some data collected on nylon-6 and a nanocomposite containing the polymer.

Studies on polypropylene nanocomposites demonstrate that thermal stabilization is achieved without increasing the heat of combustion, and without an increase in the amount of soot or carbon monoxide produced. As an added advantage, the silicate fillers used in these materials are inexpensive, and can be regenerated and recycled after char formation.

A simple explanation for these effects is that the multilayered clay component within the nanocomposites act as both thermal insulators and barriers to mass transport, therefore hindering the release of volatile polymer decomposition products and enhancing the fire retardancy. Importantly, the benefits conferred by the nanocomposites structure do not appear to degrade the required strength or flexibility of the material compared with the polymer.

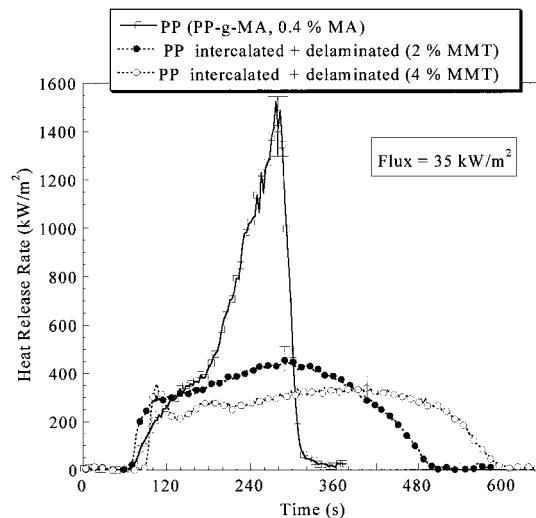


FIGURE 6 Comparison of heat release rates for nylon-6 and nylon-6-clay nanocomposites. (Reprinted from Gilman, J. (1999). "Flammability and thermal stability studies of polymer layered-silicate (clay) nanocomposites," *Appl. Clay Sci.* **15**, 31–49, with permission from Elsevier Science.)

C. Structural/Mechanical Properties

Many structural polymers, including nylons, polypropylene, polystyrene, fluoropolymers, and epoxies have been combined with layered silicates to form nanocomposites that display improved mechanical stability compared to the pristine polymers or fiber-reinforced polymer micro/macromposites. Mechanical properties that show significant improvement for nanocomposites have included tensile and flexural strengths, flexural modulus, and heat distortion temperature (HDT). In **Fig. 7**, the

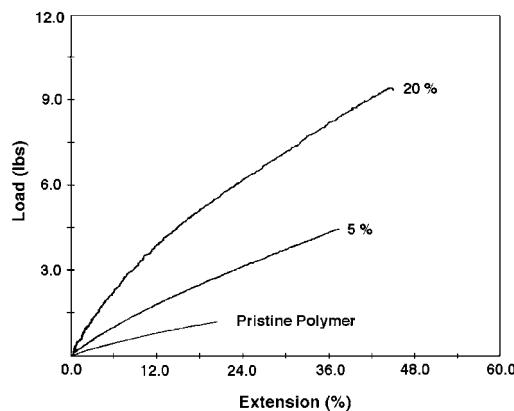


FIGURE 7 Stress-strain curves for a pristine epoxy elastomer and epoxy-clay nanocomposites prepared from epoxy with 5 and 20 wt% of an organoclay derivative of montmorillonite. The load cell had a capacity of 20 lb, and the strain rate was 0.06 in./min. (Reprinted from Lebaron, P., Wang, Z., and Pinnavaia, T. J. (1999). "Polymer-layered silicate nanocomposites," *Appl. Clay Sci.* **15**, 11–29, with permission from Elsevier Science.)

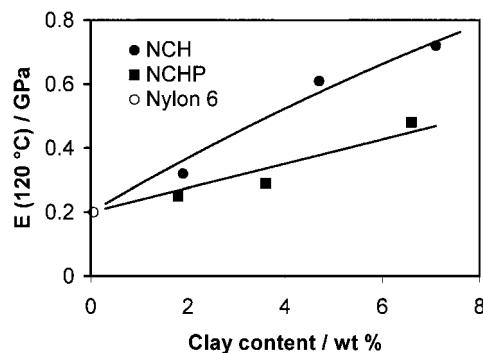


FIGURE 8 Dependence of tensile modulus, E, at 120°C on clay content for nanocomposites with nylon-6 and montmorillonite (NCH) or saponite (NCHP). The value for the polymer itself can be seen at 0% clay content. (Reprinted with permission from Kojima, Y., Usuki, A., Kawasumi, M., Okada, A., Fukushima, Y., Kurauchi, T., and Kamigaito, O. (1993). *J. Mater. Res.* **8**, 1185.)

stress-strain curves for a pristine epoxy elastomer and epoxy-clay nanocomposites prepared from epoxy with 5 and 20 wt% of an organoclay derivative of montmorillonite show the increased modulus for the nanocomposites. Figure 8 shows the dependence of tensile modulus (E) at 120°C on clay content for nanocomposites containing nylon-6 and montmorillonite (NCH) or saponite (NCHP). The value for the polymer itself can be seen at 0% clay content. Table IV illustrates the differences between the mechanical and other properties of some nanocomposites as compared with the polymers themselves. Some other technological advantages have also been reported for nanocomposites in structural applications, including the formation of materials with minimum shrinkage and warpage without elevating the densities of the products.

Significant progress has been made in the commercial applications of nanocomposites as light and tough automobile interior and exterior parts, aircraft interiors,

TABLE IV Physical and Mechanical Properties of Nylon-Clay Nanocomposites

Properties	Nylon 6	Nylon 6-clay nanocomposites	Clay wt%
Tensile strength (MPa)	68.6	97.2	4.7
Tensile modulus (GPa)	1.11	1.87	4.7
Flexural strength (MPa)	89.3	143	4.7
Flexural modulus (GPa)	1.94	4.34	4.7
Izod impact strength (J/m)	20.6	18.1	4.7
Charpy impact strength (KJ/m ²)	6.21	6.06	4.7
Heat distortion temperature 1.82 MPa (°C)	65	152	4.7
Coefficient thermal expansion (x,y)	6.3×10^{-5}	13×10^{-5}	4
Water absorption (%)	0.87	0.51	4

lawnmowers, and electrical/electronic housings. In the early 1990s, Toyota Central Research laboratories developed nylon-clay based nanocomposites technology and now has a broad-based proprietary patent technology that is licensed to companies in Japan and the United States. Ube Industries, as an example, produces timing belt covers from nylon-clay nanocomposites for automotive applications. Various grades of nylons and engineering plastic-based nanocomposites are now commercially available. In addition to fundamental research in universities, both American and Japanese industries are actively developing products for automotive applications.

Nanocomposites derived from thermoplastic and thermosetting polymers can be processed by conventional technologies such as extrusion, injection molding, blow molding, and thermoforming. During processing, the mechanical and rheological properties are likely to change, and it is important to understand these changes. The addition of fillers generally results in an increase in viscosity in conventional micro- or macrocomposites materials. Nanocomposites, on the other hand, exhibit shear thinning at higher shearing rate, so that viscosity values can be lower than those of pure polymer. These observations suggest that nanocomposites are amenable to melt-processing methods. The shear thinning effect, although not fully understood, may be related to the reorientation of the inorganic sheets when under shear.

D. Chemical Barrier Properties for Packaging

Nanocomposites can form films or coatings that are useful gas or solvent barriers for packaging applications. The exfoliated nanocomposite structure results in an increase in the effective pathlength for molecular diffusion because the layered inorganic filler forces permeating gas and solvent molecules to travel a highly tortuous path. This effect can greatly reduce gas and moisture transmission through the film. For example, oxygen permeability through Nylon 6 is reduced by about 50% in the Nylon-6 based nanocomposites. Also, nanocomposite packaging films based on poly(ethylene terephthalate) PET have been studied as replacements for applications where conventional polymer films are currently employed. In Fig. 9 the permeability of CO₂ through polyimide and nanocomposite films is indicated for different clay contents. As indicated in the side illustration, the path for gas diffusion becomes highly tortuous when clay sheets are added, and the use of higher aspect ratio clays increases the effect.

E. Optical Properties

Unlike microcomposites that have particle sizes in range of several millimeters, nanocomposite films that show minimal light scattering can be prepared for some

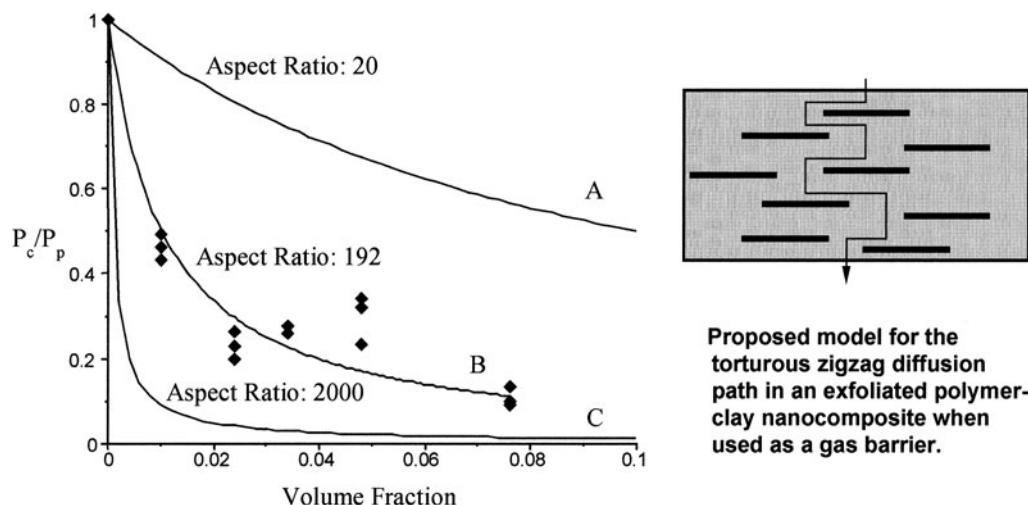


FIGURE 9 Relative permeability of CO_2 through polyimide and nanocomposite films as a function of clay content. The use of high aspect ratio clays results in a more dramatic decrease in CO_2 transmission through the film. (Reprinted with permission from *Chem. Mater.* (1994), **6**, 573–595, American Chemical Society.)

host/intercalate combinations. Film coatings of this type are therefore potentially useful in optical applications. As an example, the refractive indices of some transparent polymers such as poly(acrylic acid), poly(vinyl alcohol), poly(acrylamide), and PMMA are similar to those of inorganic glasses. Therefore, coatings of such polymers, which may improve chemical or mechanical stabilities of the product, can be applied without loss of optical clarity. The formation of nanocomposites allows the selection of a range of refractive indices as well as a further improvement in resistance properties. Such optical coatings are now in development for optical lenses, dispersive optical elements, band light filters, and nonlinear optical materials. One important additional function lies in blocking ultraviolet light. Figure 10 demonstrates the de-

creased uv transmission for nylon films when a clay filler is added.

Often, molecular electronics and photonic devices require materials that display a combination of electrical, magnetic, and/or optical properties. Clearly, it would be difficult to find single compounds that display all the appropriate properties. However, multifunctional materials such as nanocomposites can sometimes be prepared by design. For example, donor–acceptor substituted stilbazoliums constitute a class of organic materials with very high second-order nonlinear optical (NLO) response. Layered metal phosphorous trichalcogenide, MPS_3 ($\text{M} = \text{Fe}$, Mn , and Ni) possesses interesting magnetic properties. A nanocomposite combining magnetic and NLO properties has been prepared by intercalating dimethylamino-N-methyl stilbazolium cations into the MnPS_3 structure.

Poly(phenylene vinylene) (PPV) displays useful optical properties for electroluminescence and nonlinear optical conversion. Nanocomposites containing the polymer appear to have greater thermal and chemical stabilities, which may be very significant for successful application in devices. Another potentially useful effect is the polymer ordering (or poling) induced by the host lattice. Studies have concluded that PPV composite structures in silica and V_2O_5 can be used in waveguides.

A nanocomposite derived from nylon-11 polymer and inorganic solid, BiI_3 , exhibits significant X-ray photoconductivity by synergistically combining properties of the parent phases. The high X-ray absorption efficiency of the heavy-element (Bi) component is enhanced by the desirable low dark conductivity, dielectric properties, and processability due to the polymeric component. This type of material is of interest in the development of diagnostic

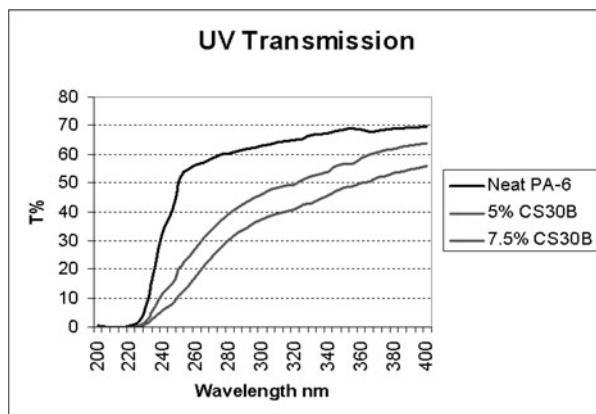


FIGURE 10 UV transmission comparison for nylon and nylon-clay nanocomposite films with 5 and 7.5 wt.% filler. Reprinted with permission from Southern Clay Products.

digital radiographic tools such as magnetic resonance imaging (MRI) and computer-aided tomography (CAT).

F. The Future

Based on basic research results over the past few years, nanocomposites have been proposed for numerous applications, including rechargeable battery electrodes, electrocatalysts, photoconductors, MRI, and high-frequency dielectrics. Certainly, as the understanding of synthesis and process control improves, conventional composites or polymers will be increasingly replaced by this new generation of materials.

Despite the recent focus by the scientific community, it must be noted that polymers have long been included within nanostructured assemblies, because such materials arise in biological systems. Biosynthetic nanocomposites include shells, bones, and cobwebs. These are complex materials derived from biopolymers (lipids, proteins, and carbohydrates) and inorganic or biosynthetic minerals (such as calcium carbonate and calcium phosphate). A widely investigated and complex example is the nacre in abalone shell, which owes its resilience to a multilayer, nanoscaled laminate of biopolymers between microdimensional platelets of the aragonite form of calcium carbonate. This assembly leads to remarkable mechanical properties, and the desire to synthesize analogs generates much research interest. Without a doubt, future studies will also improve our understanding and ability to recreate and improve upon these biological materials.

ACKNOWLEDGMENTS

The authors thank Professors Thomas Pinnavaia, Allan Jacobson, Emmanuel Giannelis, and Harmut Fischer, and Drs. Anton Lerf, Arimitsu

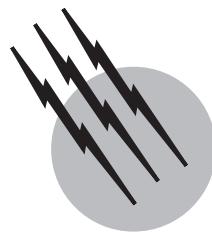
Usuki, Douglas Hunter, and Jeffrey Gilman for providing figures and source materials. MML gratefully acknowledges the National Science Foundation for grants supporting basic research on layered nanocomposites and intercalation chemistry.

SEE ALSO THE FOLLOWING ARTICLES

BATTERIES • BIOPOLYMERS • COMPOSITE MATERIALS • ELECTROLYTE SOLUTIONS, THERMODYNAMICS • NANOELECTRONICS • NANOSTRUCTURED MATERIALS, CHEMISTRY OF • POLYMERS, ELECTRONIC PROPERTIES • POLYMERS, MECHANICAL BEHAVIOR • POLYMERS, THERMALLY STABLE • SOL-GEL PROCESSING

BIBLIOGRAPHY

- Alberti, G., and Bein, T. (eds.) (1996). "Comprehensive Supramolecular Chemistry," Elsevier, Oxford.
- Herron, N. (1997). Synthesis, Characterization, and Properties of Semiconductor Nanoclusters. In "Handbook of Nanophase Materials" (A. Goldstein, ed.) Marcel Dekker, New York.
- Klein, L. C. (1997). Sol Gel Formation and Deposition. In "Handbook of Nanophase Materials" (A. Goldstein, ed.) Marcel Dekker, New York.
- LeBaron, P., Wang, Z., and Pinnavaia, T. (1999). "Polymer-layered silicate nanocomposites: An overview," *Appl. Clay Sci.* **15**, 11–29.
- Lerf, A. (2000). Intercalation compounds in layered host lattices: Supramolecular chemistry in nanodimensions. In "Handbook of Nanostructured Materials and Nanotechnology" (H. Nalwa, ed.), Vol. 5, pp. 1–166, Academic Press, New York.
- Lerner, M., and Oriakhi, C. (1997). Polymers in Ordered Nanocomposites. In "Handbook of Nanophase Materials" (A. Goldstein, ed.), pp. 199–219, Marcel Dekker, New York.
- Oriakhi, C. (2000). "Polymer nanocomposition approach to advanced materials," *J. Chem. Ed.* **77**, 1138–1146.
- Ozin, G. A. (1992). "Nanochemistry: Synthesis in diminishing dimension," *Adv. Mater.* **4**, 612.
- Sanchez, C., Ribot, F., and Lebeau, B. (1997). "Molecular design of hybrid organic-inorganic nanocomposites synthesized via sol-gel chemistry," *J. Mater. Chem.* **9**, 35–44.



Phosphors

Alok M. Srivastava

General Electric Company

- I. Basic Phosphor Theory
- II. Fluorescent Lamp Phosphors
- III. Quantum-Splitting Phosphors (QSPs) and Mercury-Free Fluorescent Lamps
- IV. Phosphors for Plasma Display Panels (PDPs)
- V. Scintillator Materials
- VI. Cathodoluminescent Phosphors and Color Television
- VII. Phosphors for White Light-Emitting Diodes (LEDs)

GLOSSARY

Color rendering index (CRI) A unitless index, abbreviated variously as CRI or R_a , is a measure of the accuracy with which white light renders colors; by definition, a blackbody source has a CRI of 100.

Color temperature The concept of color temperature (kelvin) defines the color of a lamp—at a given temperature, the lamp color is identical to the perceived emission from a blackbody source at that temperature; lamps for general-purpose illumination are designed to fall on or in the proximity of the blackbody locus (see Fig. 3).

Lamp efficacy Lamp output per unit power input (lumen per watt or lm/W); the theoretical maximum efficiency of a light source is 683 lm/W, that being a perfectly efficient monochromatic lamp emitting at 555 nm.

Lumen The lumen is defined as 1/683 W of monochro-

matic green light at a frequency of 540×10^{12} Hz (corresponding to a wavelength of about 555 nm where the human eye is most sensitive; see Fig. 2).

Lumen equivalent Number of lumens, equivalent to 1 W of light-energy emitted by the phosphor.

Lumen maintenance This relates the change in the device brightness with time.

THE TERM LUMINESCENCE has been used to denote the emission of radiation in excess of the thermal radiation produced by heat in a given material. A material exhibiting luminescence is called as phosphor. The phosphor material is usually a crystalline inorganic solid that is deliberately doped (or activated) with a small fraction of luminescent ion. Some phosphors, such as CaWO_4 , are self-activated.

We begin with the (often-confusing) nomenclature of luminescence. *Photoluminescence* is used to denote the generation of light under ultraviolet (UV) excitation.

Cathodoluminescence is used to denote emission of radiation under cathode ray excitation. *Electroluminescence* is the generation of light from the application of electric field to a material. *Thermoluminescence* is used to denote the release of energy in the form of photons during heating of the solid that has previously absorbed energy. The energy released in the form of luminescence during chemical and biological reactions is denoted as *chemiluminescence* and *bioluminescence*, respectively. *Ionoluminescence* refers to light emission from ion bombardment. The term *mechanoluminescence* is used to denote the generation of light during any mechanical action (such as grinding) on a solid.

This article reviews the uses of phosphors that continue to be the central component of an increasing number of functional devices such as fluorescent lamps, flat panel displays, X-ray detectors, cathode ray tubes, and light-emitting diodes (LEDs). Because of space limitations, issues of the rather complex interactions of the luminescent ion with its surroundings that alter optical property, phosphor preparation procedures, and procedures employed in the determination of phosphor optical properties cannot be dealt with here in detail. Emphasis is placed on discussing the luminescence properties and limitations of specific phosphors, which currently are cited as having met the various requirements for use in the proposed device. However, the role of phosphors in creating new technologies, such as mercury-free fluorescent lamps and white light-emitting diodes, will be identified and discussed.

The background information presented in this review article is common knowledge to those associated with luminescent materials. There are a number of many useful references one might consult for general background

devices incorporating luminescent materials. A few suggestions are the following. Readers interested in comprehensive information specific to the subject presented in this article are encouraged to consult Shionoya and Yen. Strongly recommended are the monographs by Butler, by Blasse and Grabmaier, and by those edited by Kitai and by Vij.

I. BASIC PHOSPHOR THEORY

A concept that is useful for understanding much of the experimental data on the luminescence spectra of a localized ion impurity in solids is the configurational coordinate model (CCM). The CCM is illustrated in Fig. 1. The ordinate represents the potential energy of the system for the ground and the excited states. The abscissa is a configuration coordinate specifying the configuration of ions surrounding the luminescent ion. For example, the con-

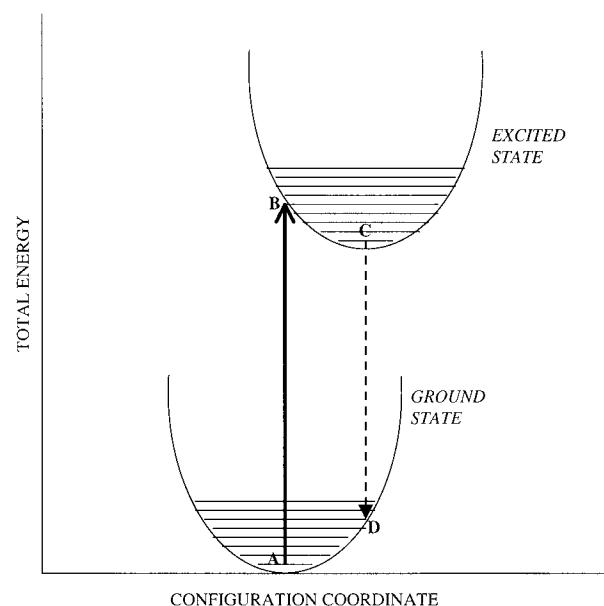


FIGURE 1 Schematic configuration coordinate diagram.

figuration coordinate may represent the distance between the luminescent ion and its nearest neighbor ions (usually anions).

We start with the assumption that the ion is in its ground electronic state and in its lowest vibrational level (Fig. 1). The center is raised to the excited state at (B) by absorption of light. The solid vertical line shows the absorption process. The Franck-Condon principle states that the time for a transition between two electronic states (of the order of 10^{-16} sec) is much faster than the vibrational period of nuclei (or crystal lattice vibrations, of the order of 10^{-13} sec). This principle permits the assumption that the absorption transition path is strictly vertical. Consequently, the atoms do not change their position during the electronic transition, and the coordinate of the center immediately following the absorption is the same as the coordinate before the absorption. For allowed transitions, the lifetime of the excited state (of the order of 10^{-8} sec) is typically much longer than the lattice vibrational period. Therefore, the absorption process is followed by the center relaxing to the lowest energy point on the excited state curve (C) by transferring its vibration energy to the lattice (emission of phonons). Luminescence can now occur from the equilibrium position of the excited state (C) to the level (D) of the ground electronic state. The system relaxes from (D) to (A) by again giving off energy to lattice vibration.

It is seen that the energy of the emitted photon is smaller than that of the absorbed photon. This fact, first recognized by the British physicist George Stokes, is called the Stokes shift. The Stokes shift that may amount to several electron volts provides a measure of the interaction of a

luminescent center with its neighboring ions. The Stokes shift is near zero for optical transitions which give rise to sharp emission lines (for example, optical transitions with the $4f^n$ levels of trivalent rare-earth ions). Under this condition, the curve representing the excited state is almost parallel to the curve representing the ground state. The Stokes shift is large for optical transitions, which give rise to broadband emission. Examples include the optical transitions emanating from the $4f^{n-1}5d^n$ excited states of ions such as Pr^{3+} , Ce^{3+} , and Eu^{2+} , charge transfer emission of niobates, tungstates, and molybdates, and emission involving the $3d$ states of transition metal ions such as Mn^{2+} . These transitions couple strongly with the lattice. For these ions, changes in the crystal structure and composition of the host lattice usually have a profound effect on the luminescence properties.

The configuration coordinate diagram can be constructed from careful measurements of the phosphor optical properties. Without going into further detail, it may be mentioned that the CCM can explain the following phenomena in luminescence:

1. The Gaussian shape of the excitation and emission bands
2. The dependence of the emission wavelength on temperature
3. The temperature broadening of the excitation and emission bands
4. The decrease in the luminescence efficiency at elevated temperature

Phosphors are usually characterized by their absolute quantum efficiency under excitation by monochromatic light of various wavelengths. The quantum efficiency is defined as the ratio of the number of photons emitted by the phosphor to the number of photons absorbed. The procedure for measuring absolute quantum efficiency can be quite difficult, and the interested reader is encouraged to consult Shionoya and Yen for further details.

Luminescence that involves energy transfer is of great interest because of the success in turning such systems into commercially viable phosphors. In such systems the incident photons are absorbed by one center (sensitizer) which then transfers their excitation to another center (activator), which in turn emits the desired radiation. The sensitizer-activator concept is utilized in cases where there is near absence of excitation by the incident primary photons when only the activator is present in the host lattice. There are various mechanisms associated with energy transfer between the sensitizer and the activator, but we will not discuss this topic here; readers interested in this topic are encouraged to consult the works by Shionoya and Yen, Butler, Blasse and Grabmaier, and Kitai. Throughout this

article we will encounter examples where the concept of sensitized luminescence has been used to generate very efficient phosphors.

II. FLUORESCENT LAMP PHOSPHORS

The glossary contains terms that are used in the lamp and display industries and are of value in understanding much of the functional device characteristics data. Fluorescent lamps are known for their high efficacy, 80–100 lm/W⁻¹, and are available in several color temperatures, depending on the choice of phosphor (3000, 3500, and 4100 K are typical; some market segments prefer high color temperatures comparable to sunlight, 5000–6500 K). As we will show, the color rendering of fluorescent lamps ranges from ~50 to the mid 90s.

Fluorescent lamp phosphors convert the UV emission of a rare-gas/mercury discharge plasma into visible (white) light. Typical fluorescent lamp discharges emit UV radiation predominately on the mercury atomic intercombinational line ($^3P_1 \rightarrow ^1S_0$) at 254 nm, but also emit radiation on the atomic mercury resonance line ($^1P_1 \rightarrow ^1S_0$) at 185 nm (depending upon conditions, 10–20% of the UV radiation is emitted at 185 nm). The discharge conversion efficiency of electric power to 254 nm UV line is high (65%). The phosphor is responsible for nearly all the visible light produced by the lamp, with the visible mercury lines contributing only a few percent to the total lamp light output. The phosphors have been optimized for efficient operation under excitation by 254 nm UV light. In this article we will focus on phosphors for general-purpose illumination (fluorescent lighting systems used in homes, offices, factories, and shops); phosphors used in specialized

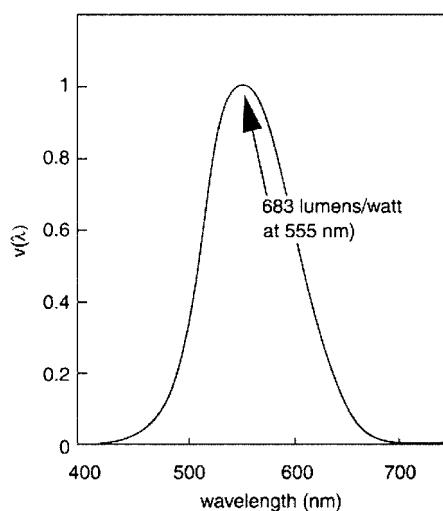


FIGURE 2 Relative sensitivity of the human eye.

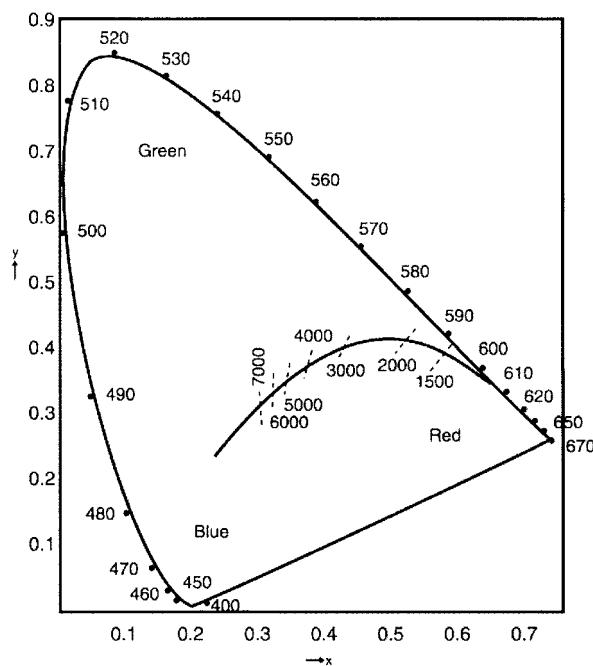


FIGURE 3 Chromaticity space; the area under the enclosed curve corresponds to physically allowed colors; BBL is the black-body locus.

lighting systems (such as sun-tanning lamps, black lights, and plant growth lamps) will be identified and discussed briefly.

A. Halophosphate Phosphor

The early fluorescent lamps used various combinations of naturally fluorescing minerals such as willemite (Mn^{2+} -activated Zn_2SiO_4) to generate white light. A significant breakthrough in fluorescent lighting occurred in the 1940s with the development of the calcium halophosphate phosphor ($\text{Sb}^{3+}, \text{Mn}^{2+}$ -activated $\text{Ca}_5(\text{PO}_4)_3(\text{Cl}, \text{F})$). As shown in Fig. 4, the complementary blue and yellow emission bands of this phosphor offer a resultant white emission in fluorescent lamps. The blue band is due to the activator Sb^{3+} ions, which absorb the 254 nm radiation of the discharge and emit a part of this energy in a band peaking near 480 nm. The excitation energy is also transferred from Sb^{3+} to Mn^{2+} resulting in the orange-red Mn^{2+} emission peaking at about 580 nm. The Sb^{3+} acts as the sensitizer and the Mn^{2+} as the activator. The ratio of the blue to orange emission can be adjusted; increasing Mn^{2+} content, for example, suppresses the blue emission and enhances the orange emission (see Fig. 4). A range of whitish color from near blue to orange can therefore be attained from a single material. A further variation in color can be achieved by changing the F:Cl ratio. Generally, the increase in Cl levels results

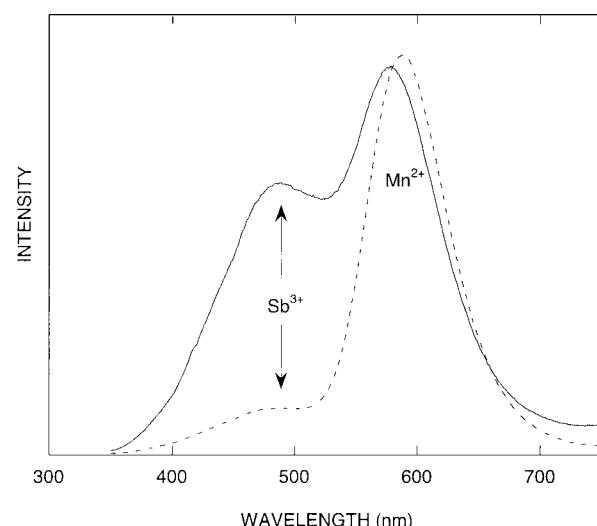


FIGURE 4 Emission spectrum of halophosphate phosphor activated with various amounts of trivalent antimony and bivalent manganese ions.

in the peak position of the Mn^{2+} emission band shifting toward the orange. The synthesis of the halophosphate phosphor is rather complex. It may be re-emphasized that we will not be discussing the phosphor manufacturing processes. In general, procedures employed in phosphor manufacturing are proprietary and are still very much an art; the reader is referred to other sources for descriptions of some of the procedures that have been devised for commercial phosphors.

The color rendering index (CRI) of “cool white” fluorescent lamps is in the upper 50s or lower 60s. The rather low CRI follows from the emission spectrum which shows that the two complementary emission bands do not fill the visible region of the spectrum and in particular are deficient in the red spectral region (Fig. 4). Hence, colors are distorted under these lamps compared to their appearance under blackbody radiation sources or sunlight. An improved CRI is obtained by blending halophosphate phosphor with a red-emitting phosphor (typically Sn^{2+} -activated strontium orthophosphate). These “deluxe” lamps display higher CRI (85) due to the continuous emission across the visible region from the phosphor blend. The improvement in CRI is, however, accompanied by a marked drop in the lamp brightness (efficacy of 50 lm/W) since the emission of red phosphor occurs in the outlying region of the visible spectrum where the eye is insensitive. This tradeoff between the luminosity and the color rendition limits the usefulness of this outstanding phosphor in many practical applications.

The color of the halophosphate lamp, although acceptable in many applications, is not sufficiently pleasing for many people to install such lamps in the living areas of

their home; hence they tend to be relegated to the basement or the garage. Much-improved light is possible through the use of modern rare-earth phosphors where white light is made by mixing red-, green-, and blue-emitting phosphors. The rare-earth phosphors and their impact on general lighting are discussed below.

B. Rare-Earth Phosphors in Fluorescent Lamps

The 1970s saw the development of a new generation of low-pressure mercury fluorescent lamps where one could simultaneously combine markedly high color rendering ($CRI \sim 80$) with high efficacy. Theoretical modeling had already predicted that a spectrum composed of emission in narrow bands near 450 nm (blue), 550 nm (green), and 610 nm (red) contribute to high CRI and efficacy. In the mid-1970s, such fluorescent lamps were commercially available and contained the blend of three rare-earth activated phosphors (the blends are popularly known as tricolor or triphosphor blends): Eu^{3+} -activated Y_2O_3 (red emitting), Tb^{3+} -activated $\text{CeMgAl}_{11}\text{O}_{19}$ (green emitting), and Eu^{2+} -activated $\text{BaMgAl}_{10}\text{O}_{17}$ (blue emitting).

We now proceed to briefly discuss the individual rare-earth based phosphors that have played a vital role in the development of tricolor lamps.

1. Red-Emitting Phosphor (Y_2O_3 : Eu^{3+} ; YEO)

The YEO phosphor absorbs the 254 nm Hg discharge emission through a charge transfer transition involving the Eu^{3+} ion and the neighboring O^{2-} ions. This charge transfer transition peaks at about 230 nm (Fig. 5). Conse-

quently, the YEO phosphor does not absorb 254 nm radiation maximally, with plaques reflecting about 25% of the incident 254-nm radiation. The high reflectivity coupled with strong absorption of the 254-nm radiation by the green- and blue-emitting phosphors necessitates the dominance (by weight) of YEO in the triphosphor blend. The reflectivity can be decreased with an increase in the Eu concentration, but the high cost of Eu_2O_3 prohibits commercial application of such compositions. The emission spectrum of this phosphor is ideal for red color generation (Fig. 5). It consists of a dominant peak at 611 nm which corresponds to the electric dipole transition, ${}^5D_0 \rightarrow {}^7F_2$.

The quantum efficiency of YEO, the highest of all lighting phosphors, is close to unity and constant in the range of Eu concentrations from 1 to 15 mol%. The lumen maintenance of this phosphor is exceptional and exceeds that of the other lighting phosphors. The phosphor has been applied in fluorescent lamps for more than 2 decades due to its high quantum efficiency, near perfect red emission, and exceptional lumen maintenance. Research aimed at replacing this phosphor with less-expensive alternatives has not met with success.

2. Green-Emitting Phosphors

All phosphors that supply the required green emission in lamps for general-purpose illumination are based on sensitized luminescence. In all cases, the trivalent cerium ion acts as the sensitizer and the trivalent terbium ion acts as the activator. Incident UV photons are absorbed by the Ce^{3+} ions, which then transfer their excitation energy to the Tb^{3+} ions, which in turn efficiently emit green photons according to the energy level and the transition rates of the Tb^{3+} ions. The optical properties of the three green-emitting phosphors are summarized in Table I and their salient features briefly discussed below.

a. $\text{CeMgAl}_{11}\text{O}_{19}$: Tb^{3+} (CAT). The composition $\text{CeMgAl}_{11}\text{O}_{19}$ is derived from $\text{PbAl}_{12}\text{O}_{19}$ (a material with the magnetoplumbite structure) by replacing Pb^{2+}

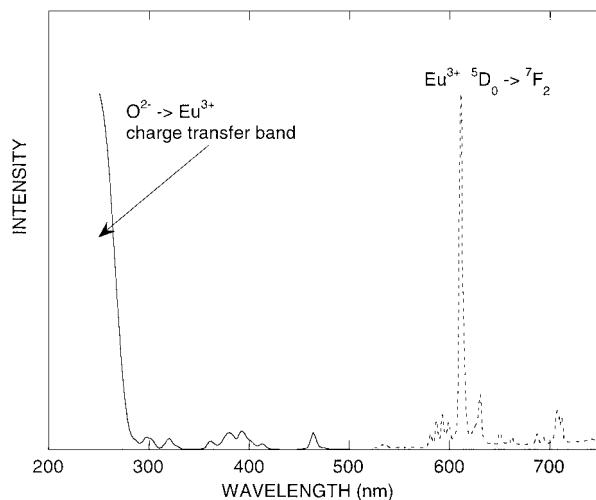


FIGURE 5 Excitation (solid line for Eu^{3+} emission at 611 nm) and emission (dotted line; excitation at 254 nm) for Y_2O_3 : Eu^{3+} .

TABLE I Optical Properties of Green Phosphors Applied in Tricolor Fluorescent Lamps

Phosphor	Emission wavelength (nm)	UV(%) ^a	VIS(%) ^a
LaPO_4 : Ce^{3+} , Tb^{3+} (LAP)	543	7	86
$\text{GdMgB}_5\text{O}_{10}$: Ce^{3+} , Tb^{3+} (CBT)	541	2	88
$\text{CeMgAl}_{11}\text{O}_{19}$: Tb^{3+} (CAT)	542	5	85

^a Quantum efficiency for ultraviolet (UV) and visible (VIS) emission.

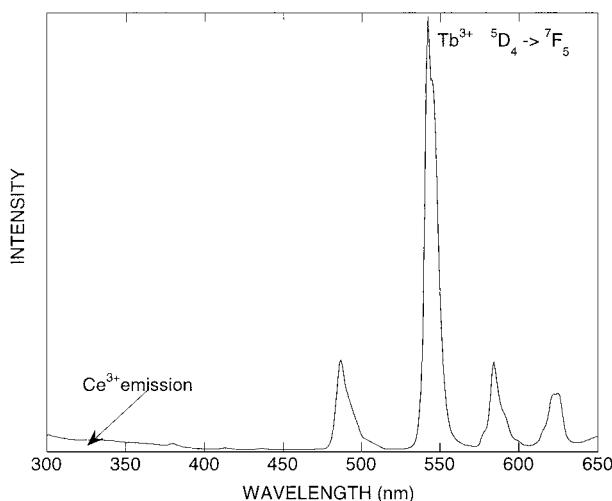


FIGURE 6 Emission spectrum for $CeMgAl_{11}O_{19}: Tb^{3+}$ (excitation at 254 nm).

by Ce^{3+} and substituting one of the Al^{3+} with Mg^{2+} . The compound $CeMgAl_{11}O_{19}$ is an efficient UV emitter when excited by 254-nm radiation. The emission arises from allowed transitions between the ground and excited states of the Ce^{3+} ion which are derived from the $4f^1$ and $5d^1$ electronic configurations, respectively.

The introduction of Tb^{3+} in $CeMgAl_{11}O_{19}$ quenches the Ce^{3+} emission and generates the green Tb^{3+} luminescence as a result of the Ce^{3+} -to- Tb^{3+} energy transfer (Fig. 6). The spectral overlap between the excitation and emission bands is small due to the relatively large Stokes shift (about 9400 cm^{-1}) of the Ce^{3+} luminescence. Hence there is an absence of energy migration among the Ce^{3+} ions in this lattice. The Ce^{3+} -to- Tb^{3+} energy transfer is limited to the six nearest neighbors at 5.6 \AA and this requires rather high Tb^{3+} concentration (33 mol%) for the complete quenching of the Ce^{3+} emission. The optimum phosphor composition ($Ce_{0.67}Tb_{0.33})MgAl_{11}O_{19}$ exhibits high quantum efficiency and excellent lumen output and maintenance during lamp operation.

b. $LaPO_4: Ce^{3+}, Tb^{3+}$ (LAP). The efficient green luminescence of $LaPO_4: Ce^{3+}, Tb^{3+}$ (monoclinic monazite-type structure) has been known since the 1970s but only recently has the material gained importance as a fluorescent lamp phosphor. The advantages of this phosphor over $CeMgAl_{11}O_{19}: Tb^{3+}$ are the manufacturing ease due to the lower synthesis temperatures (about $1000^\circ C$) and the lower Tb concentrations required for optimum performance. The LAP emission spectrum is dominated by the transition at 543 nm ($Tb^{3+}, ^5D_4 \rightarrow ^7F_5$) (Fig. 7).

The role of Ce^{3+} in LAP is that of the sensitizer. The allowed $Ce^{3+} 4f \rightarrow 5d$ transitions are in resonance

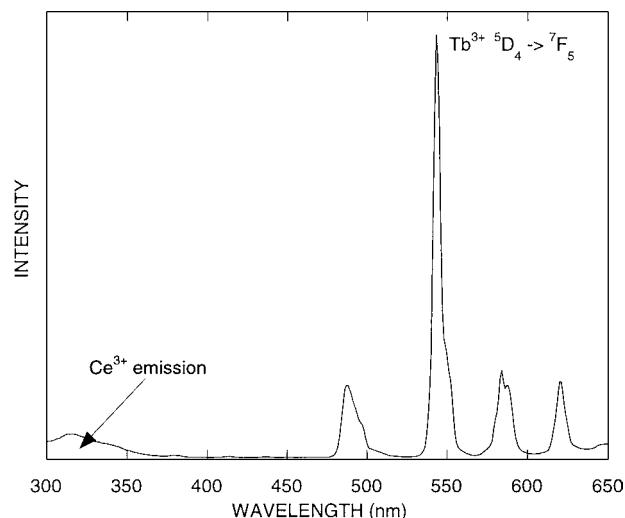


FIGURE 7 Emission spectrum for $LaPO_4: Ce^{3+}, Tb^{3+}$ (excitation at 254 nm).

with the mercury discharge. In striking contrast to CAT, considerable energy migration occurs over the Ce^{3+} ions prior to the Ce^{3+} to Tb^{3+} transfer. This reduces the amount of Tb^{3+} required for optimum performance; the commercial formulation contains 27% Ce and 13% Tb ($La_{0.60}Ce_{0.27}Tb_{0.13}PO_4$). Energy migration among the Ce^{3+} ions results from the large spectral overlap between the excitation and emission bands due to the relatively small Stokes shift (about 4700 cm^{-1}) of the Ce^{3+} luminescence in $LaPO_4$. The phosphor displays excellent lumen output and maintenance during lamp operation.

c. $GdMgB_5O_{10}: Ce^{3+}, Tb^{3+}$ (CBT). In CBT, the Gd^{3+} ions assist in the transport of energy from the sensitizer (Ce^{3+}) to the activator (Tb^{3+}) ions. The Ce^{3+} emitting levels are resonant with the lowest Gd^{3+} excited levels (6P_J). Energy transfer from Ce^{3+} to Gd^{3+} energy transfer is an efficient process. Efficient green luminescence is generated when the activator (Tb^{3+}) ion intercepts the excitation energy that is percolating over the Gd^{3+} ions (Fig. 8). The quantum efficiency is high, and the phosphor displays excellent stability in fluorescent lamps.

Substituting a part of Mg^{2+}/Zn^{2+} (present in a distorted octahedral coordination) with Mn^{2+} ions produces a deep red-emitting phosphor. The Mn^{2+} ions efficiently trap the migrating Gd^{3+} excitation energy and display a broadband emission peaking at 620 nm. The phosphor is used in high-color-rendering fluorescent lamps ($CRI \sim 95$), the so-called five-band lamps (these lamps have four or five phosphors instead of the three of the triphosphor lamp). Such fluorescent lamps with $CRI (>90)$ close to that of an incandescent lamp find application in color-critical

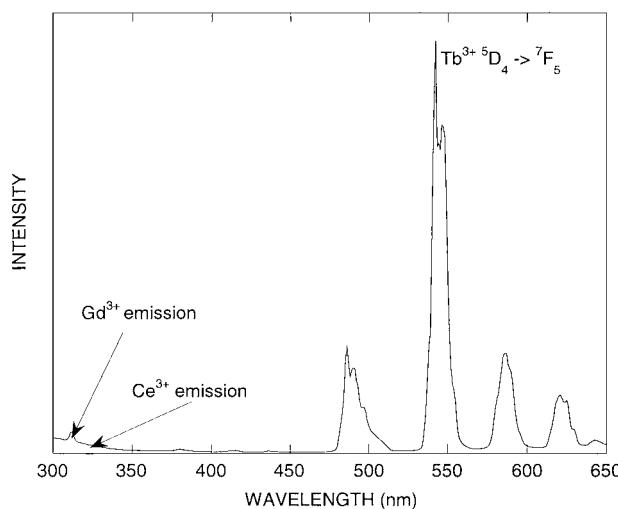


FIGURE 8 Emission spectrum for GdMgB₅O₁₀: Ce³⁺, Tb³⁺ (excitation at 254 nm).

applications such as museum lighting. The increase in CRI occurs at the expense of the lamp efficacy.

3. Blue-Emitting Phosphors

Two blue-emitting phosphors are commonly used in tri-color fluorescent lamps. One is Eu²⁺-activated BaMg Al₁₀O₁₇, a material with the beta-alumina structure. Efficient Eu²⁺ luminescence with emission maximum at 450 nm supplies the required narrow-band blue emission in the triphosphor blend. The strong absorption in the UV region is due to the allowed 4f⁷5d → 4f⁶5d¹ transitions.

The second commonly used blue phosphor is Eu²⁺-activated (Sr, Ba, Ca)₅(PO₄)₃Cl, a material with the apatite (halophosphate) structure. The phosphor displays strong UV absorption with a narrow-band emission peaking at 450 nm (Fig. 9). The blue phosphors represent only a mirror weight fraction of the triphosphor blend (about 10% for color temperature of 4100 K). However, blends designed for higher color temperatures, say 6500 K, require higher amounts of the blue-emitting component.

The great interest in phosphors based on rare-earth ions is largely due to the success in recent years in turning the luminescence of these ions into commercially important phosphors. Specifically, fluorescent lamps with high efficacy and high color rendering would not have been realized without the use of rare-earth phosphors. Further, compact fluorescent lamps (CFLs) would not have been possible without the development of rare-earth phosphors, because their higher wall temperature and higher UV flux quickly degrades halophosphate materials.

The chief drawback of rare-earth phosphors is their high cost that has more than doubled the cost of some fluorescent lamps. Discovering materials that produce optimum

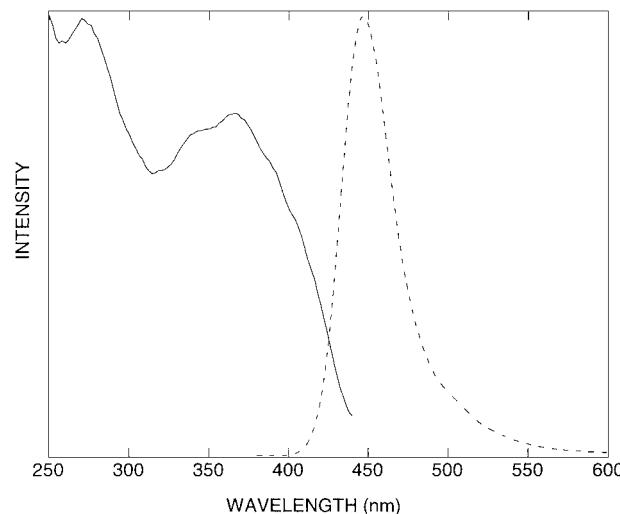


FIGURE 9 Excitation (solid line for Eu²⁺ emission 450 nm) and emission (dotted line; excitation at 254 nm) for (Sr,Ca,Ba)5(PO₄)₃Cl:Eu²⁺.

performance at reduced rare-earth ion concentration, or in some cases, by replacing the expensive rare earths by transition metal ions (such as Mn²⁺) will reduce the phosphor cost. Decreasing the phosphor particle size may also reduce cost. A covering of four layers of phosphor particles requires proportionally less material with decreasing particle size. The reduction in particle size is limited by the UV absorption strength of the individual phosphor particle. There is also considerable interest in the lighting industry in decreasing the loss of phosphor efficiency during lamp operation.

C. Phosphors for High-Pressure Mercury Vapor Lamps

The high-pressure mercury vapor (HPMV) discharge (pressure exceeding several atmospheres; efficiency of 20–25%) produces visible emission that is particularly deficient in the red spectral region. The discharge also produces significant amounts of longer wavelength 365-nm UV radiation. Since power is dissipated in a small bulb, the operating temperature of the HPMV lamp (HPMVL) exceeds 200°C. For safety reasons, the small bulb in which the discharge is confined is enclosed in a larger outer bulb. The color rendering of HPMVL can be improved by a phosphor capable of efficiently converting the short- and long-wavelength UV to red emission. Hence, illumination based on HPMVLs requires phosphors capable of red emission at high temperature (200–250°C).

From the standpoint of the pertinent requirements, three phosphors have been proposed for this application: Mn⁴⁺-activated magnesium fluorogermanate, Eu³⁺-activated Y(P,V)O₄, and Sn²⁺-activated Sr-Mg orthophosphate.

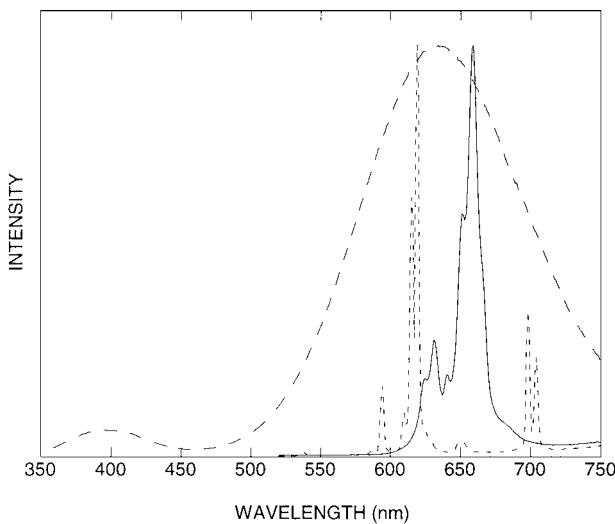


FIGURE 10 Emission spectrum for $(\text{Sr},\text{Mg})_3(\text{PO}_4)_2:\text{Sn}^{2+}$ (dashed line), $\text{Mg}_4\text{GeO}_{5.5}\text{F}:\text{Mn}^{4+}$ (solid line), and $\text{YVO}_4:\text{Eu}^{3+}$ (dotted line); excitation at 254 nm.

The emission spectrum of the three phosphors is shown in Fig. 10.

The emission of Mn^{4+} -activated magnesium fluorogermanate consists of narrow bands between 600 and 700 nm. The emission is far in the red that enhances the CRI of HP-MVL but makes a poor match to the eye sensitivity curve. This coupled with the fact that the phosphor absorbs significantly throughout the visible (the body color of the phosphor is yellow) leads to poor lamp lumen output.

The activation of $(\text{Sr}, \text{Mg})_3(\text{PO}_4)_2$ by Sn^{2+} results in a broadband emission centered at 630 nm. As the phosphor does not absorb visible radiation, both the CRI and lumen output is increased.

The emission spectrum of Eu^{3+} -activated $\text{Y}(\text{P}, \text{V})\text{O}_4$ is dominated by strong lines in the proximity of 620 nm. The UV radiation is absorbed by the vanadate group, and the energy is transferred to Eu^{3+} ions. Unlike the other two phosphors, the line emission of Eu^{3+} does not occur in the outlying region of the visual spectrum where the eye is insensitive (see Fig. 2). Consequently, the lumen equivalent of this phosphor is high. The limitation of this phosphor is in the rather poor absorption of the long wavelength UV radiation generated by the discharge. With respect to the effect of variation in temperature, all three phosphors show relatively good thermal stability of their emission intensity.

D. Phosphor for Specialized Illumination Lamps

We now give a brief account of the phosphors that find application in specialized illumination. Generally, in spe-

cialized illumination we encounter reasons for the precise yet subtle control over the phosphor spectral power distribution. Consider the example of the UV-emitting sun-tanning lamp. Sun tanning is accomplished by exposing the human skin to radiation in the UV-A spectral region (320–400 nm) which results in fast tanning and darkening of the skin, and to radiation in the UV-B spectral region (260–320 nm) which promotes the long-lasting reddening of the skin. It is particularly important that the spectral output of the sun-tanning lamp comprise a balanced output of UV-A and UV-B radiation. During the past 2 decades, our understanding of the effects of UV radiation on the human skin has increased, and this has led to changes in the spectral output of sun-tanning lamps. Currently, the phosphor $\text{Ba}_2\text{SiO}_5:\text{Pb}^{2+}$ (broadband emission centered at 350 nm) in a blend with $\text{SrAl}_{12}\text{O}_{19}:\text{Ce}^{3+}$ (broadband emission centered at 300 nm) is used to deliver a balanced amount of UV-A and UV-B for skin tanning. The phosphor Eu^{2+} -activated SrB_4O_7 has been proposed as an intense source of UV-B radiation and has been used in the past in sun-tanning lamps. The borate and silicate phosphors further find use in mineral lamps (black lights).

Lamps based on the $\text{Gd}^{3+} \ ^6\text{P}_{7/2} \rightarrow \ ^8\text{P}_{7/2}$ sharp line emission are used to control psoriasis, a skin disorder that cannot be cured but can be controlled by UV radiation. The Gd^{3+} ion does not satisfactorily absorb incident UV light at 254 nm and hence requires sensitization. Currently, two phosphors satisfy the pertinent requirements for this lamp type: Pr^{3+} -activated GdB_3O_6 and Bi^{3+} -activated $(\text{Gd},\text{La})\text{B}_3\text{O}_6$. In these phosphors, the Pr^{3+} and Bi^{3+} ions sensitize the Gd^{3+} emission.

The efficient broadband emission of Eu^{2+} in $\text{Sr}_2\text{P}_2\text{O}_7$ that is centered at 420 nm is used in photocopy lamps. Your local grocery shop uses red-enhanced lamps based on the emission of Mn^{4+} -activated magnesium fluorogermanate to enhance the appearance of red meat (see Fig. 10). This phosphor is also used in specialty lamps for enhanced plant growth (photosynthesis).

III. QUANTUM-SPLITTING PHOSPHORS (QSPs) AND MERCURY-FREE FLUORESCENT LAMPS

Conventional fluorescent lamps which provide energy-efficient general-purpose lighting in commercial and residential applications employ mercury as the active species for generation of the UV radiation. There is increasing concern, however, with the leaching of soluble mercury from spent lamps in solid-waste landfills entering the groundwater supplies. A fluorescent lamp where low-pressure xenon discharge excites suitable phosphors to generate white light can be envisioned as a

mercury-free replacement for existing fluorescent lamps. Recently, xenon gas discharge efficacy of nearly 65% has been demonstrated under optimum operating conditions. However, lamp efficiency issues prevent us from any considerations of conventional phosphors as white light-generating materials in such a fluorescent lamp.

The overall conversion efficiency of a fluorescent lamp can be written schematically as $\eta_{\text{lamp}} \sim \eta_{\text{uv}} [\varepsilon_{\text{vis}}/\varepsilon_{\text{uv}}] Q_p$, where η_{uv} is the discharge efficiency for converting electric power to UV power, Q_p is the quantum efficiency of phosphor, ε_{vis} is the weighted average energy of the spectrum of visible photons emitted by the phosphor (this is fixed by the spectral sensitivity of the human eye, which peaks near 555 nm), and ε_{uv} is the energy of the photon emitted by the discharge and absorbed by the phosphor.

For conventional mercury-based fluorescent lamps, the efficiency is (very approximately): 0.25~0.65 [254 nm/555 nm] 0.85. Note that the discharge efficiency is about two-thirds, and the phosphor converts nearly every incident photon into UV radiation. If the discharge is 65% efficient, and if the phosphor is nearly perfect, what accounts for the relatively low overall conversion efficiency of 25%? The answer is in the Stokes shift, denoted here by the ratio $[\varepsilon_{\text{vis}}/\varepsilon_{\text{uv}}]$ which accounts for the fact that each UV photon incident on the phosphor carries nearly 5 eV of energy, while each photon emitted by the phosphor carries barely over 2 eV. This single process accounts for 55% of energy loss in the conventional fluorescent lamp.

If we wish to reproduce the energy conversion efficiency of conventional fluorescent lamps but with a Xe discharge that emits at 147 nm, the higher Stokes shift loss can be offset by higher phosphor quantum efficiency. There has been some demonstration of phosphors which, on average, produce more than one visible photon for each incident UV photon. We refer to such materials as “quantum-splitting phosphors”(QSPs). For example, the phosphor YF₃:Pr³⁺ yields a room temperature quantum efficiency of 1.40 ± 0.15 under excitation by 185 nm radiation. If this phosphor also yields the same quantum efficiency under 147 nm excitation, then the energy conversion requirement becomes more reasonable: 0.25~0.65 [147 nm/555 nm] 1.40. One can immediately recognize the benefits of the YF₃:Pr³⁺ phosphor in devices that employ vacuum ultraviolet (VUV) emission of rare gas discharge as the primary exciting source.

The process of quantum splitting in Pr³⁺-activated phosphors is shown in Fig. 11A. Incident VUV photons are absorbed via an allowed Pr³⁺ $4f \rightarrow 5d$ optical transition. The excitation decays to the 1S_0 level. The transition probability is then such that the 1S_0 level decays radiatively to the 1I_6 level resulting in the generation of the first photon. A second transition that connects the upper

3P level with several of the ground-state levels yields the second photon.

It is unfortunate that the practical uses of YF₃:Pr³⁺ phosphor is not straightforward, for several reasons. First, the phosphor is not stable in the presence of rare-gas/mercury discharge as is used in conventional fluorescent lamps. Whether this instability arises because of chemical, photochemical, plasma-assisted, or other mechanism is not known. Second, large-scale manufacture of fluorinated materials is difficult. Third, the Pr³⁺ emission that occurs mainly in the deep blue (at about 405 nm) is essentially wasted because the human eye is virtually insensitive to that wavelength.

The aforementioned problems with the practical implementation of fluorinated materials led Srivastava and colleagues to pursue the development of optimized oxide host lattices as QSPs. Three oxide materials in which Pr³⁺ is observed to quantum split were discovered: SrAl₁₂O₁₉, LaMgB₅O₁₀, and LaB₃O₆. However, none of the oxide materials displayed quantum efficiency exceeding unity, and the problem of the deep blue emission still remained.

Recently, an effort in QSPs centered on the trivalent gadolinium ion has been described in the literature. Incident VUV photons are absorbed via the Gd³⁺ $^8S_{7/2} \rightarrow ^6G_J$ optical transition (Fig. 11B). A cross-relaxation process excites the emission of the intentionally added activator, Eu³⁺ (Step 1 in Fig. 11B). During this cross-relaxation process the Gd³⁺ ion relaxes to the lower 6P_J state. Energy migrating over the 6P_J levels is trapped by a second Eu³⁺ ion (Step 2 in Fig. 11B). Hence, two red photons may be produced per incident VUV photon. Indeed, internal quantum efficiencies approaching nearly two in the Li(Y,Gd)F₄:Eu³⁺ have been estimated.

The foregoing discussion shows that the phosphor that appeared as a weak link in the energy conversion chain can be improved by developing QSPs. No such material has been turned into a commercially viable phosphor, although significant efforts continue in the development of such phosphors.

IV. PHOSPHORS FOR PLASMA DISPLAY PANELS (PDPs)

In recent years several companies have made sizable investments for mass producing color plasma display devices. A color plasma flat-panel display consists of an intermittent atmosphere-pressure xenon discharge, which excites red, green, and blue phosphors to form color images. The PDPs essentially consist of top and bottom glass plates. This structural simplicity makes the fabrication of large area display (40–60 inch diagonal) possible with a thickness of 10 cm or less. In a PDP, phosphors convert

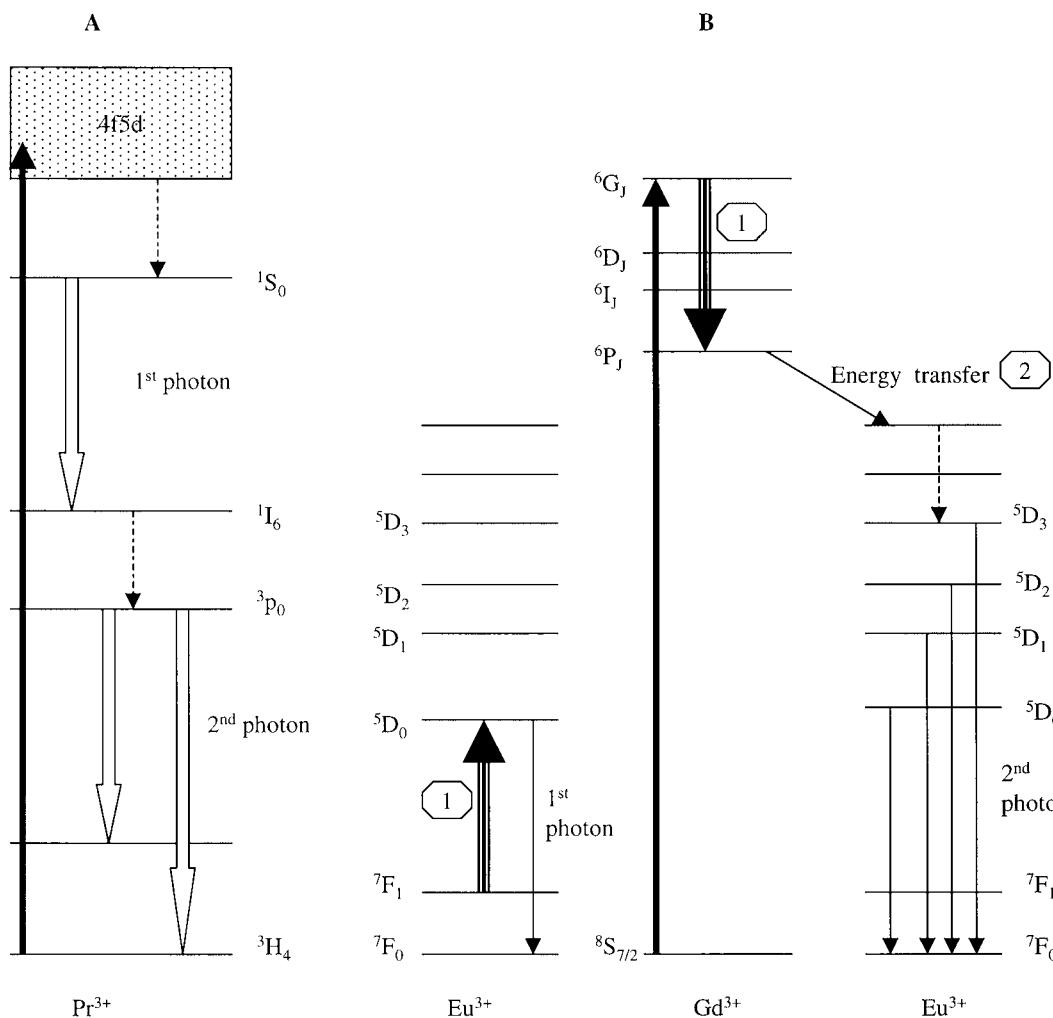


FIGURE 11 Schematic illustration of quantum splitting in (A) Pr^{3+} -activated materials and (B) $\text{Gd}^{3+}, \text{Eu}^{3+}$ -activated materials; ----> indicates nonradiative transitions.

the VUV emission of xenon discharge plasma into visible light. The gas discharge is controlled individually in each cell of a PDP. The desirable 147 and 173 nm VUV radiations are produced when the electronically excited Xe monomers and excimers, respectively, decay radiatively to the ground state. Relative to the mercury discharge in low-pressure mercury fluorescent lamps, the discharge conversion efficiency of electric power to VUV radiation in PDPs is rather inefficient, about 6%. The inefficiency is related to the cell geometry.

Divalent europium-activated $\text{BaMgAl}_{10}\text{O}_{17}$ which supplies the required narrow blue emission in triphosphor fluorescent lamps can also be applied in PDPs since the phosphor is efficiently excited to a blue emission by VUV irradiation. The limitation of this phosphor in PDPs is traced to instability under irradiation by VUV light.

The Mn^{2+} -activated Zn_2SiO_4 (Fig. 12) and $\text{BaAl}_{12}\text{O}_{19}$ phosphors are efficiently excited to a saturated green luminescence by VUV photons and are currently the best available phosphor for this application. To reduce tailing in moving images, phosphors with short decay times are desirable. The spin forbidden $^4T_1 \rightarrow ^6A$ optical transition, responsible for the green emission in these materials, exhibits a rather long decay time (tens of milliseconds). This limits the use of Mn^{2+} -activated phosphors in displays where the image changes rapidly. Interestingly enough, an increasing proportion (to a certain extent) of Mn^{2+} in Zn_2SiO_4 decreases the lifetime of the excited state while having no deleterious effect on the quantum efficiency. Hence, a tradeoff between efficiency and decay time may yield phosphor compositions that are optimal for use in moving image display devices. Although efficient green-emitting phosphors based on the activator Tb^{3+} are

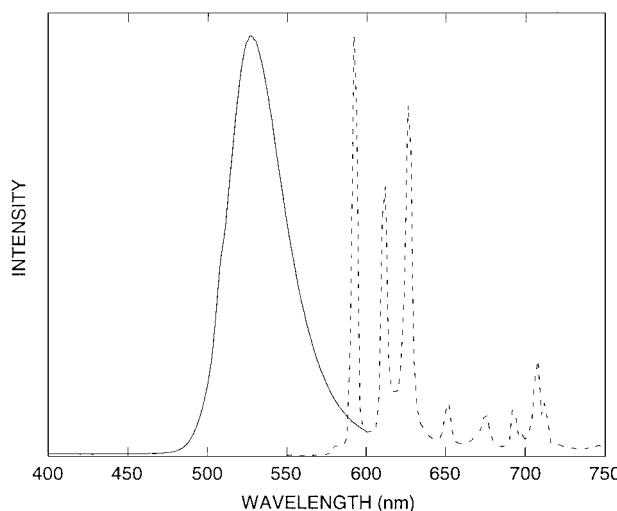


FIGURE 12 Emission spectra of $\text{Zn}_2\text{SiO}_4:\text{Mn}^{2+}$ (solid line) and $(\text{Y},\text{Gd})\text{BO}_3:\text{Eu}^{3+}$ (dotted line) phosphors applied in plasma display devices.

known, their color of emission (mainly consisting of a sharp line emission near 545 nm) is not suitable for use in PDPs.

From the experimental data that has been presented on the variation in the emission characteristics of Eu^{3+} within the $(\text{Y}_{1-x}\text{Gd}_x)\text{BO}_3$ system of phosphors, the composition $(\text{Y}_{0.65}\text{Gd}_{0.35})\text{BO}_3:\text{Eu}^{3+}$ has been chosen to supply the primary red color in PDPs (Fig. 12).

Let us consider the phosphor energy conversion efficiency in a PDP. Nearly 70% of the input power will be lost when the phosphor converts each 7 eV (172 nm) Xe photon to a visible photon with an average energy of about 2 eV of energy (the phosphor is only about 30% efficient in terms of energy). Further, nearly 75% of the energy is lost in the conversion of an incident 147 nm xenon photon into a ~ 555 nm visible photon. Clearly, any development of efficient QSPs (see above) would be of great interest to

manufacturers of color PDPs. The low overall efficiency (about 1 lm/W for white light; in contrast, the efficacy of conventional cathodray tube (CRT) is about 3 lm/W for white light) and lumen depreciation are both significant problems in current PDPs.

V. SCINTILLATOR MATERIALS

High-density luminescent materials that convert X-rays, and various other high-energy particles (such as α -, β -, and γ -rays) directly to visible light are denoted as scintillators. The need to stop the highly penetrating primary radiation in short distances requires that scintillators have high density and contain a large proportion of elements with high atomic number (high Z elements). Although scintillators may be in the form of glasses, liquids, and gases, this review focuses on inorganic solids that are commercially useful as scintillators. The field of scintillators is quite extensive, as these materials find increasing scientific and commercial use in such diverse fields as medical imaging, industrial inspections, X-ray security apparatus, and high-energy physics calorimetry, to name a few. Such phosphors may be used, also, to convert high-energy radiation into radiation which actuates photosensitive devices such as photodiodes, charge-coupled devices (CCDs), and photo-multiplier tubes (PMTs) whose output offers an efficient means of converting high-energy radiation into electrical signal. In the following discussion, only certain outstanding examples of currently known scintillators and their uses are cited. Relevant properties of some scintillators have been given in Table II. The physical process in scintillators may be classified into three parts: (i) the absorption of high-energy primary photons by the host and the conversion into thermalized electrons and holes (conversion efficiency, β), (ii) transfer of the electrons and holes produced to the luminescent center (transfer efficiency, S), and (iii) luminescence process (quantum efficiency, Q).

TABLE II Properties of Some Scintillator Materials

Scintillator	Application	Density (g/cm ⁻³)	Decay time (second)	Emission wavelength (nm)
CsI:Tl^+	Radiology	4.51	980×10^{-9}	550
$\text{Gd}_2\text{O}_2\text{S:Tb}^{3+}$	Radiology	7.34	3×10^{-3}	540
BaFBr:Eu^{2+}	Radiology	5.20	800×10^{-9}	390
$(\text{Y},\text{Gd})_2\text{O}_3:\text{Eu}^{3+}$	CT	5.95	1×10^{-3}	611
$\text{Gd}_2\text{O}_2\text{S:Pr}^{3+}$	CT	7.34	3×10^{-6}	513
CdWO_4	CT	7.99	14×10^{-6}	480
NaI:Tl^+	Nuclear camera	3.67	230×10^{-9}	415
$\text{Bi}_4\text{Ge}_3\text{O}_{12}$	PET	7.13	300×10^{-9}	480
$\text{Gd}_2\text{SiO}_5:\text{Ce}^{3+}$	PET	6.71	40×10^{-9}	430

Hence the energy efficiency η of a scintillator may be written as

$$\eta = \beta * S * Q \quad (1)$$

The energy of the incident high-energy primary photon will generally be much greater than the band gap (E_g) of the material. Consequently, more than one luminescence photon is excited through the very large number of electron–hole pairs created within the phosphor lattice by the absorption of the incident primary photons with energy E_p . In terms of energy efficiency, however, the performance of the scintillator is less impressive, and for a photon of energy E_l (which is generally less than E_g) generated in the scintillating process, this efficiency is given by

$$\eta = (E_l/\alpha E_g) S * Q \quad (2)$$

where α is host lattice dependent and is typically greater than 2. Consider the case where S and Q are unity in a scintillator material with E_l approaching E_g . In this case the energy efficiency should be about 25 to 30% which is in agreement with the reported efficiencies with the best phosphor materials.

A. Phosphors for Intensification of X-Ray Produced Images

Among the numerous useful applications of phosphors are the very important scintillator phosphors for medical imaging that enable life-saving diagnostics tools. In one medical imaging modality the use of scintillators has been in the intensification of X-ray produced images, where the phosphor screens are placed adjacent a photographic emulsion and the screen/film combination exposed to X-rays. The patient is placed between X-ray source and the phosphor, screen/film cartridge. During medical diagnoses, it is essential to keep the X-ray dosage to the patient as low as possible. The phosphor converts each incoming X-ray photon to hundreds or thousands of visible photons that form the image on the film. In lieu of the phosphor, an unacceptable high exposure of X-rays to the patient would be required to obtain a good image on the film itself (more than 99% of X-ray passes through the film and is essentially lost with only a small fraction of it being absorbed by the film emulsion). Consequently, the chief use of the phosphor is to reduce the X-ray dosage to the patient.

The thickness of the phosphor screen determines the sharpness (resolution) of the X-ray image. Thin screens with dense packing of the phosphor powder increase film spatial resolution. Consequently, the phosphor particle size and morphology play a significant role in determining the optimum coating densities and hence the film resolution. Phosphors with short persistence are required in

X-ray screens, in order to minimize blurring or fogging which results in a vestigial (ghost) image of one patient onto another on subsequent film exposure. Electronic defects in solids that capture the electron and holes created in the conduction and valence bands are mainly responsible for the afterglow. It is thus not surprising that afterglow is mainly controlled by the phosphor synthesis technique.

The scheelite CaWO_4 was one of the first luminescent materials to be proposed in 1896 (one year after the discovery of X-rays by Roentgen) as a scintillator for the efficient conversion of X-ray photons to visible (blue) light. Calcium tungstate is found in nature as the mineral scheelite. The large calcium ions have eight oxygen neighbors forming a distorted cube, whereas the smaller tungsten ions occur in tetrahedral coordination. The blue luminescence of this material is due to charge transfer transition occurring between the W^{6+} ions and the coordinating O^{2-} ligands. The charge transfer transitions are thus localized within the WO_4^{2-} groups. The disadvantages of this phosphor are the relatively low density (6.06 g/cm^{-3}) and X-ray absorptivity in the 30- to 80-KeV energy range, its rather poor X-ray-to-visible photon conversion efficiency (5%), and the rather long persistence. Among the new and improved scintillators which have been devised for application with blue and green sensitive photographic emulsions are the blue-emitting LaOBr:Tm^{3+} , green-emitting $\text{Gd}_2\text{O}_2\text{S:Tb}^{3+}$, and UV-emitting YTaO_4 . These are commercially useful intensifier scintillators for X-ray photons in the 15- to 100-KeV energy range due to their high X-ray-to-light conversion efficiencies, high density, and minimum afterglow. The salient features of the aforementioned phosphors are now briefly discussed.

1. LaOBr:Tm^{3+}

The efficient intraconfigurational transitions of the Tm^{3+} ($4f^{12}$) ion in the UV and blue spectral region coupled with the high density (6.13 g/cm^{-3}) and efficient X-ray-to-light conversion efficiency (18%) are the chief reasons for the popularity of this material as an X-ray intensifying screen phosphor. Due to the layered structure of LaOBr (isostructural with BaFCl , another important intensifying screen phosphor), particle morphology of the phosphor prepared by flux-assisted solid state is plate-like, which packs rather poorly in a screen and exhibits the tendency for light piping toward the side of the plate. The phosphor is also water sensitive, although this problem has been resolved through the addition of antimony tartrate in the final product. Despite its limitations, the phosphor has achieved great practical significance due to its meeting the various requirements for use in X-ray screens.

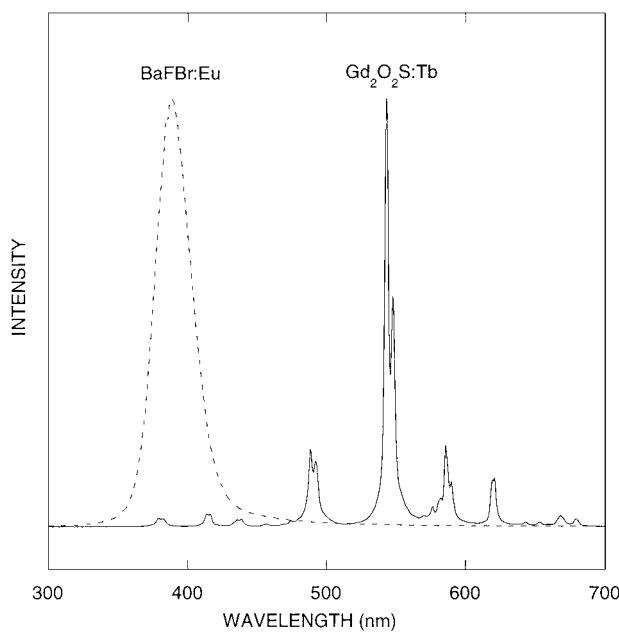


FIGURE 13 X-Ray (60 KVp) excited emission spectra of scintillators used in radiology; BaFBr:Eu²⁺ (solid line) and Gd₂O₂S:Tb³⁺ (dotted line).

2. Gd₂O₂S:Tb³⁺

The high density, strong X-ray absorption at the Gd K-edge (50 KeV, in the middle of the diagnostic X-ray energy range), high conversion efficiency (19%), and efficient luminescence due to the intraconfigurational transitions of the Tb³⁺ ($4f^{12}$) ion in the blue and green spectral region (depending on the Tb concentration; Fig. 13) coupled with the crystallization of the phosphor in perfect polyhedra results in the oxy sulfide as one of the most commercially important X-ray phosphors. Recently, Gd₂O₂S:Pr³⁺ has been proposed as a scintillator for use in computed tomography (CT; to be discussed below).

3. YTaO₄

Among the new X-ray phosphors, one of the most efficient and stable occurs in one of the structural modifications of YTaO₄. Reports in the archival literature show the occurrence of three crystal forms of this material: a high temperature form with the tetragonal (T) scheelite structure (previously discussed) which distorts to a monoclinic (M) form. The M-modification is isostructural with the mineral fergusonite. Yet another monoclinic modification, denoted as M'-YTaO₄ in the literature, is synthesized below about 1400° C. It is this modification that is efficiently excited by X-rays to UV luminescence (emission is a broadband with peak maximum occurring at about 330 nm). As in CaWO₄, the luminescence is due to charge trans-

fer transitions within the tantalate groups. Some members of this versatile general family of phosphors such as M'-LuTaO₄ (density=9.75g/cm⁻³) and M'-GdTaO₄ offer higher densities and X-ray absorption but at significantly higher materials cost (especially M'-LuTaO₄). Starting with M'-YTaO₄ and adding an increasing proportion of Nb activator up to $x = 0.020$ in M'-Y(Ta_{1-x}Nb_x)O₄ results in a new broad emission band appearing at a longer wavelength (emission maximum at 410 nm) due to charge transfer transition within the niobate groups.

B. Storage Phosphors

It is instructive to examine a technology introduced in 1983 in medical radiography where a phosphor containing an appreciable density of traps (the trap depths, ΔE , are large compared to kT) is able to store the free charge carriers (electrons and holes) created by X-ray irradiation. At a later time, the stored energy can be released by the absorption of a just-sufficient amount of energy to overcome the energy barrier (ΔE). The radiative recombination energy of the freed charge carriers is converted into a luminescence photon. The stimulation can be thermal or optical; luminescence generated by thermal and optical stimulation is denoted as thermally stimulated luminescence (TLS) and photostimulated luminescence (PSL), respectively.

The PSL phosphor used in nearly all commercial digital X-ray imaging system is BaFBr:Eu²⁺. Upon X-ray irradiation, the created holes are trapped by the Eu²⁺ ions resulting in Eu³⁺ while the electrons are trapped by halide ion vacancies resulting in the formation of F centers. Illumination in the F-center absorption band by a He-Ne laser (red light at 633 nm) provides a sufficient amount of red-photon energy to stimulate recombination of the trapped electron with the holes trapped at the Eu²⁺ site. This results in a broadband emission at 390 nm corresponding to the Eu²⁺ $4f^65d \rightarrow 4f^7$ electronic transition (Fig. 13). Consequently, the latent image in the screen is read out by scanning the screen with a focused red light source. We should mention that from the available evidence in the archival literature it appears that the origin of PSL in BaFBr:Eu²⁺ is rather complex and still controversial.

C. Computed Tomography

In *computed tomography* (CT), the X-ray source and opposed detector rotate 360° in a plane around the patient. The imaging device measures the attenuation of X-rays through the body and provides excellent cross sectional images of the human body. Three scintillators that are currently used in commercial CT detectors are CdWO₄, (Y,Gd)₂O₃:Eu³⁺, and Gd₂O₂S:Pr³⁺.

We have already discussed the luminescence of $\text{Y}_2\text{O}_3:\text{Eu}^{3+}$ (see Section II.B.1). The addition of Gd^{3+} is necessary to increase the physical density of the phosphor; at the current density of 5.95 g/cm^{-3} , the material is able to stop 99.96% of the 70-keV X-ray in 3 mm. The high melting point of this material creates problems in developing single crystal, and so ceramists have developed techniques of fabricating these materials as polycrystalline ceramics that are sintered to transparency. The red emission is very well suited for detection by photodiodes. The problem of afterglow in this material is largely overcome by the incorporation of ions such as Pr and Tb in $(\text{Y},\text{Gd})_2\text{O}_3:\text{Eu}^{3+}$. The material is also radiation resistant. Such custom-made phosphors satisfy nearly all requirements of generating visible light with high efficiency, reduced afterglow, and excellent stopping power while displaying high resistance to damage under ionizing radiation.

The luminescence of CdWO_4 (wolframite structure) in the blue-green (broad emission band centered at 480 nm) is a result of charge transfer transition occurring within the tetrahedral tungstate groups (Fig. 14). The rather low melting point of 1240°C permits the attainment of large single crystals of this material. The material is susceptible to radiation damage; the defects created by X-ray irradiation are capable of absorbing the visible luminescence of the scintillator. Refinements in the synthesis procedures have resulted in a minimization of impurity defects and reduction in Cd vacancy defects.

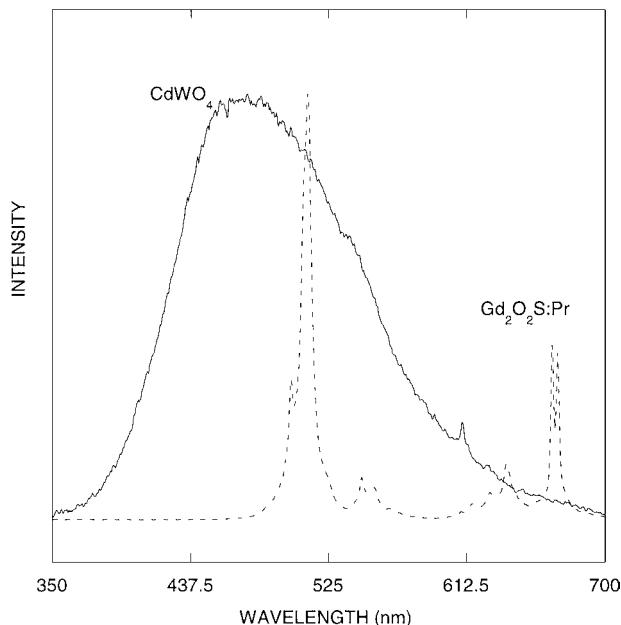


FIGURE 14 X-ray (60 KVp) excited emission spectra of scintillators used in computed tomography; CdWO_4 (solid line) and $\text{Gd}_2\text{O}_2\text{S:Pr}^{3+}$ (dotted line).

The efficient green luminescence of $\text{Gd}_2\text{O}_2\text{S:Pr}^{3+}$ is due to the intraconfigurational transitions of the $\text{Pr}^{3+}(4f^2)$ corresponding to the ${}^3\text{P}_0 \rightarrow {}^3\text{H}_J, {}^3\text{F}_J$ electronic transitions (Fig. 14). Addition of Ce to this material reduces afterglow. The high physical density and the short decay time (about $3 \mu \text{ sec}$) resulting from the spin allowed optical transitions of the activator ion makes the phosphor attractive for use in CT detectors.

D. Nuclear Cameras and Positron Emission Tomography (PET)

Finally, we briefly discuss the luminescent materials that are utilized in *nuclear cameras* and PET. These medical diagnostic techniques are based on the detection of high-energy γ -rays emitted from a patient that has ingested short-lived radiopharmaceuticals. The spatial distribution of these pharmaceuticals yields information on the biological activity of the internal organ of interest. In nuclear cameras the scintillator detector is directly coupled with a PMT. In PET the radiopharmaceuticals generate positrons whose rapid annihilation results in the generation of two 512-keV γ -rays that are emitted in opposite directions. A ring of detectors surrounds the patient, and timing coincidence is used to detect an annihilation event. In PET it is essential to have scintillators with decay times (radiative decay rates of less than 300 nanoseconds) and high density ($>7 \text{ g/cm}^{-3}$).

Large single crystals of the popular scintillator, NaI:Tl^{+} , are easily grown by the Bridgeman-Stockbarger technique for application in nuclear cameras. The material is efficiently excited to a blue luminescence by γ -rays. The high light yield and short decay time (230 nsec) are adequate for X-ray photon counting devices. The emission corresponding to the $\text{Tl}^{+} {}^3\text{P}_1 \rightarrow {}^1\text{S}_0$ optical transition is well matched to the spectral response of high efficiency blue sensitive PMTs (Fig. 15). The scintillator is hygroscopic and must be protected from ambient atmosphere to avoid decomposition and deterioration of the material. The reported efficiency of the scintillator is about 12%. Since $E_g = 5.9 \text{ eV}$, $E_l = 2.98 \text{ eV}$ (about 415 nm), Eq. (2) suggests that S and Q must be near unity.

Large single crystals of the high-density (7.13 g/cm^{-3}) scintillator $\text{Bi}_4\text{Ge}_3\text{O}_{12}$ (BGO) are grown by the Czochralski technique for applications in PET scanners and calorimetry in high-energy physics. Although the decay time is short (about 300 nsec) and afterglow weak, its light output is relatively low compared with NaI:Tl^{+} . The low luminescence efficiency of this material at room temperature is due to the very large Stokes shift ($17,600 \text{ cm}^{-1}$). The energy efficiency of BGO amounts to only 2% due to the lower quantum efficiency of

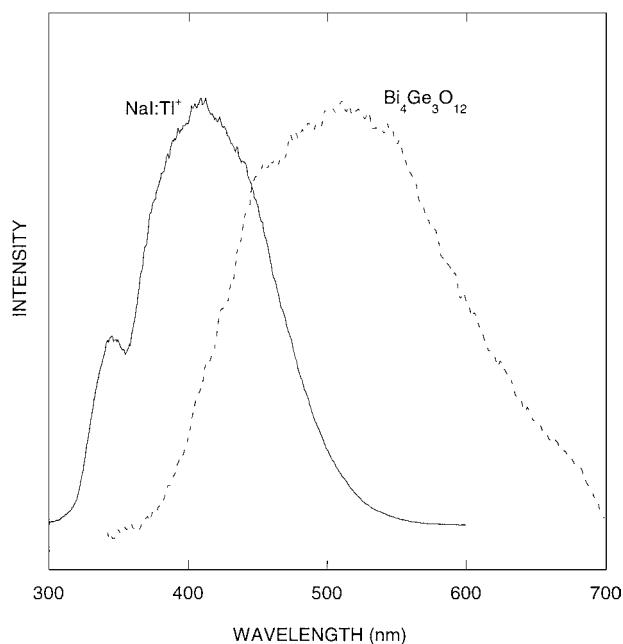


FIGURE 15 X-ray (60 KVp) excited emission spectra of scintillators used in nuclear cameras and positron emission tomography; NaI:Tl⁺ (solid line) and Bi₄Ge₃O₁₂ (dotted line)

luminescence. The emission maximum peaking at 480 nm couples well to high-efficiency PMTs (Fig. 15).

Insofar as decay times alone are concerned, phosphors based on the allowed electric dipole emission transition of the Ce³⁺ ion provide for the development of scintillators with extremely fast decay. It is instructive to examine the scintillating properties of phosphors based on rare-earth orthosilicates (Ln₂SiO₅, where Ln = Gd and Lu) with the added activator, trivalent cerium. The scintillating efficiency of Lu₂SiO₅:Ce³⁺ (density = 7.4 g/cm⁻³) approaches that of NaI:Tl⁺. The decay time is of the order of 40 nsec and the emission due to the allowed Ce³⁺ 5d → 4f optical transition peaks at 420 nm. Evidence in the literature suggests that the light output is strongly dependent on the growth and annealing conditions. The scintillating efficiency of Gd₂SiO₅:Ce³⁺ (density = 6.71 g/cm⁻³) is lower than that of Lu₂SiO₅:Ce³⁺. The decay times are short (30–60 nsec) with the emission peaking at 440 nm. Large crystals of the scintillators can be grown by the Czochalski technique.

VI. CATHODOLUMINESCENT PHOSPHORS AND COLOR TELEVISION

The direct conversion of cathode rays into visible light by phosphors makes possible major commercially important

practical devices such as color television, oscilloscopes, and electron microscopes, to name a few. Today, the annual world's production of CRTs exceeds one hundred million units. In a color TV picture tube the colored images are produced by electron beams from three separate guns impinging from slightly different angles onto a triad arrangement of phosphor stripes, each of which emits one of the primary colors: red, green, or blue. The width of the stripes is so small (less than 300 μm; there are more than one million for each of the primary colors) that they are not resolved at normal viewing distances; instead the viewer has the impression of a single blended color. The phosphors are excited with a focused electron beam; each phosphor pixel is excited for a very short time (during 1000 hr of device operation each pixel has seen the electron beam for only 10 sec); and the power input into a screen element is very high. On this basis alone, the phosphors must be optimized for operations under high excitation densities. The phosphor must also be resistant to loss of brightness during 20,000 hr of operating lifetime. In addition to brightness, color TV requires that colors are well reproduced and therefore well controlled.

Relative to the direct view CRTs, the phosphors are subjected to yet higher excitation densities in projection television (PTV) where images from three monochrome CRTs (red, green, and blue) are optically projected by a lens system onto a projection screen. The need for higher illumination on larger screens requires the higher excitation densities (>10 mJ/cm² per pulse) in PTV. However, it is not possible to make effective use of very high excitation densities since the variation of the phosphor cathodoluminescence efficiency as a function of increasing excitation density is not linear so that saturation of luminescence occurs. Further, the high excitation densities increase the phosphor temperatures in PVT to near 100°C. Consequently, phosphors that exhibit higher efficiencies at low excitation densities (as in direct view CRTs) may not necessarily be applicable in PTV. Two problems encountered at higher operating temperatures are the decrease in phosphor luminescence efficiency and the shift from white balance (ratio of red to green to blue to obtain white light on the screen). The salient features of the current phosphors that are applied in CRT and PTV are briefly presented below.

Phosphors based on ZnS exemplify the parent phosphor system from which have developed the present blue- and green-emitting CRT phosphors, (Zn, Cd)S:Cu, Al, Ag. The standard color TV blue, ZnS:Ag, Cl (or Al) has a broad emission spectrum peaking at 445 nm. Its high CRT radiant efficiency probably approaches the theoretical limit. The luminescence is of the donor–acceptor recombination type. The Ag⁺ acts as the acceptor while the Cl⁻ is a

shallow donor. The standard color TV green, ZnS:Cu,Cl (or Al) has a broad emission band peaking at 530 nm. The acceptor states (Cu) are situated at higher energies relative to the Ag acceptor states within the forbidden gap of the host lattice. The donor–acceptor recombination thus results in emission at longer wavelengths.

For the primary red, various phosphors have been proposed. The red-emitting Mn²⁺-activated phosphate, Zn₃(PO₄)₂, was the original red color TV phosphor. The red-emitting phosphor (Zn,Cd)S:Ag was introduced in 1960 when RCA introduced the all-sulfide phosphor screen. Although the efficiency of this phosphor is near the theoretical limit of cathode ray efficiency, the broad emission band makes a poor match to the eye sensitivity curve resulting in low lumen equivalent. In 1964 the rare-earth-activated phosphor, YVO₄:Eu³⁺ (see Fig. 10), the brightness of which exceeds that of the Mn²⁺-activated Zn₃(PO₄)₂ phosphor, was proposed for use in color TV. The emission of the vanadate phosphor consists of a single sharp line peaking at 620 nm which matches well with the eye sensitivity curve relative to the other red photon generating phosphors. However, the vanadate phosphor has been largely superseded for color TV by two other Eu³⁺-activated phosphors: Y₂O₃:Eu³⁺ (see Fig. 5) and Y₂O₂S:Eu³⁺. Both phosphors have slightly higher conversion efficiencies than the vanadate and their slightly more orange emission results in higher lumen equivalent.

We had previously specified the pertinent phosphor requirements for application in PTV. The green-emitting ZnS:Cu,Cl phosphor that is applied in direct view CRTs is limited by strong saturation and the steady decrease of brightness with increasing temperature. The green-emitting Gd₂O₃:Tb³⁺ phosphor (see Fig. 13) is widely used despite the rather poor temperature dependence of the luminescence efficiency (relative to room temperature, a brightness drop of 10% at 60°C and 35% at 100°C is observed). Green-emitting phosphors based on members of the garnet family, Y₃(Al,Ga)₅O₁₂, with added activator Tb³⁺ provide efficient cathodoluminescent phosphors for application in PTV. Other potential PTV green-emitting phosphors based on the luminescence of Tb³⁺ ion are InBO₃:Tb³⁺ and Y₂SiO₅:Tb³⁺. Another very efficient CRT green-emitting phosphor is LaOBr:Tb³⁺. The problems associated with the phosphor are the plate-like morphology, susceptibility to attack by water, and degradation of luminescence efficiency during operation. Literature reports suggest that these problems have been overcome, making the oxybromide phosphor a viable candidate for application in PTV.

The need for a standard color PTV blue is acute, and one struggles to identify good candidates. The blue-emitting phosphor, ZnS:Ag⁺, remains the best available

phosphor for this particular application despite its strong saturation at high excitation densities. Other phosphors that have been proposed include (La,Gd)OBr:Ce³⁺ and La(Ga,Al)O₃:Tm³⁺ (limited by saturation effects; the phosphor competes with ZnS: Ag⁺ only at excitation densities that are well above the operating levels of present PTV).

VII. PHOSPHORS FOR WHITE LIGHT-EMITTING DIODES (LEDs)

Light emitting diodes are based on semiconductors which emit spontaneous radiation under suitable forward bias conditions. The physics of LED devices is not dealt with here, and the reader is referred to an article by DenBaars in the work by Kitai. The recent burgeoning interest in the field of LEDs is due to the breakthrough by NICHIA in the epitaxial growth of InGaN compound semiconductor material which has lead to the development of the blue light-emitting diode (and blue laser diode). The semiconductor is an efficient blue light emitter with emission maximum near 450 to 470 nm and external quantum efficiency of about 10%. As in the fluorescent lamp scheme of generating white light, the blue emission of the LED may be combined with a red- and green-emitting phosphor to generate white light (Fig. 16). The high absorption of the blue LED light and the efficient conversion into photons of the desired wavelength are the principle determinants of the phosphor ability to generate white light in the blue LED + phosphor combination. Green-yellow-emitting phosphors based on members of the garnet family, (Y,Gd)₃(Al,Ga)₅O₁₂, with added activator Ce³⁺ meet most requirements for this application. As shown in Fig. 16 the lowest energy allowed 4f → 5d excitation transition of the trivalent cerium ion in Y₃Al₅O₁₂ (YAG:Ce) is in resonance with the blue “discharge” of the LED. Hence the incident LED blue radiation is absorbed strongly by YAG:Ce. The efficient emission of YAG:Ce is a broadband centered at about 580 nm (Fig. 16). The complementary blue (from the LED) and yellow (from the phosphor) emission bands offer a resultant white emission (analogous to the halophosphate phosphor, previously discussed). By careful control of the chemical composition of the garnet phosphor type, the Ce³⁺ emission can be tuned from 510 to 590 nm. Consequently, white light-emitting solid-state devices with color temperature ranging from 3000 K to greater than 6000 K can be produced.

We had earlier alluded to the phosphor energy conversion efficiencies in fluorescent lamps (roughly 40% with primary 254 nm UV) and PDPs (roughly 30% with primary 172 nm VUV). The development of a high-efficiency source of 450 nm blue radiation is a considerable

White Light

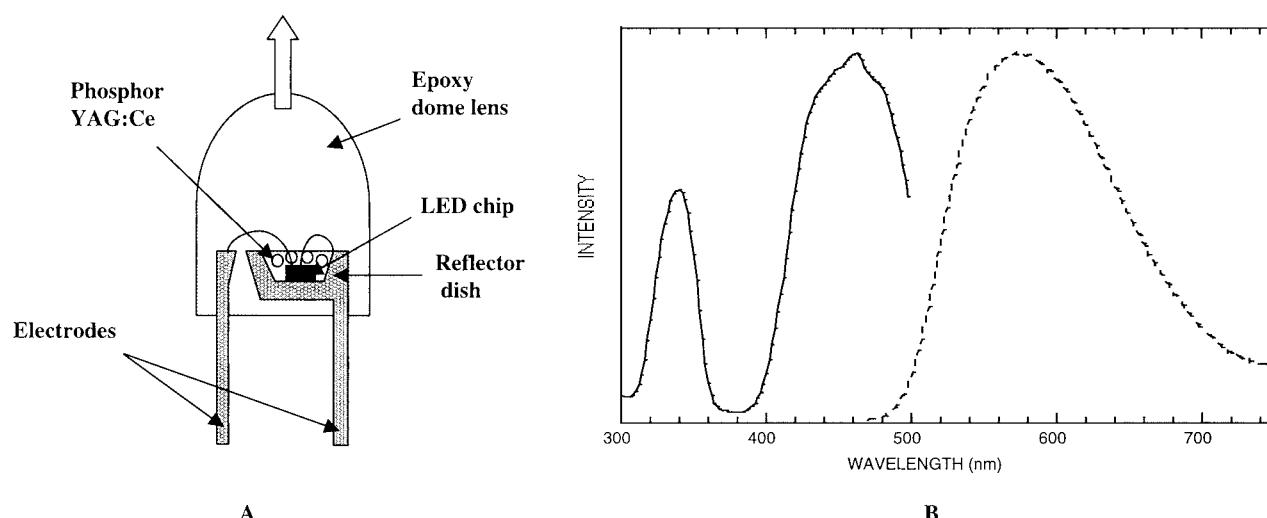


FIGURE 16 (A) Principle of white light generation in LED; (B) Excitation (solid line) and emission spectra (dotted line) of $\text{Y}_3\text{Al}_5\text{O}_{12}:\text{Ce}^{3+}$.

step forward in general illumination since the phosphor-assisted conversion of blue light energy into visible light is now about 80% efficient.

The development of a 370-nm UV LED device is possible by changing the indium content in GaN. The development of white light-emitting devices is also possible by using this source to excite a blend of three phosphors, each of which emits one of the three primary colors: red, green, and blue. Phosphors based on the narrow line emission of trivalent rare-earth ions like Eu^{3+} (red) and Tb^{3+} (green) can supply red and green components, and in combination with a suitable blue-emitting phosphor can produce white light with high efficacy. It is, of course, axiomatic that the phosphors must absorb the 370-nm long-wave UV. Fluorescent lamp phosphors have been optimized for efficient operation under excitation by 254 nm UV. Hence, it is necessary to custom make new phosphors to meet the requirements of this application.

The generation of white light by the UV LED/phosphor technology has one very desirable property (just like in the fluorescent lamp technology and unlike blue LED/phosphor technology): the electrical-to-UV radiation conversion process performed by the semiconductor is largely independent of the color of the emitted visible light by the phosphor. This affords very precise control on the color of the resultant white light so that it is possible to develop LEDs with different spectral output (for general and specialized applications) without disturbing the semiconductor technology. Consequently, one can separately optimize the semiconductor and phosphor properties for a particular application. We note that the

phosphor energy conversion in the UV LED/phosphor technology is about 65% efficient.

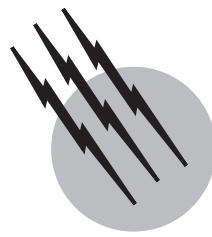
SEE ALSO THE FOLLOWING ARTICLES

IMAGING OPTICS • ION KINETICS AND ENERGETICS • LIGHT EMITTING DIODES (LEDs) • LIGHT SOURCES • RADIATION EFFECTS IN ELECTRONIC MATERIALS AND DEVICES • X-RAY ANALYSIS

BIBLIOGRAPHY

- Blasse, G. (1994). *Chem. Mater.* **6**, 1465.
- Blasse, G., and Grabmaier, B. C. (1994). "Luminescent Materials," Springer-Verlag, Berlin/New York.
- Bourcet, J. C., and Fong, F. K. (1974). *J. Chem. Phys.* **60**, 34.
- Butler, K. H. (1980). "Fluorescent Lamp Phosphors," Pennsylvania State Univ. Press, University Park, PA.
- Duclos, S. J. (1998). *Interface* **7**, 34.
- Issler, S. L., and Torardi, C. C. (1995). *J. Alloys Comp.* **229**, 54.
- Justel, T., Nikol, H., and Ronda, C. R. (1998). *Angew Chem. Int. Ed.* **37**, 3084.
- Kitai, A. H. (1993). "Solid State Luminescence; Theory, Materials and Devices," Chapman and Hall, London.
- Koedam, M., and Opstelten, J. J. (1971). *Lighting Res. Tech.* **3**, 205.
- Piper, W. W., Ham, F., and DeLuca, J. A. (1974). *J. Lumin.* **8**, 341.
- Ronda, C. R., and Amrein, T. (1996). *J. Lumin.* **69**, 245.
- Shionoya, S., and Yen, W. M. (1999). "Phosphor Handbook," CRC Press, Boca Raton, FL.
- Smets, B. M. J. (1987). *Mater. Chem. Phys.* **16**, 283.
- Sommerdijk, J. L., and Versteegen, J. M. P. J. (1974). *J. Lumin.* **9**, 415.
- Sonoda, M., et al. (1983). *Radiology* **148**, 833.
- Srivastava, A. M., and Beers, W. W. (1997). *J. Lumin.* **71**, 285.

- Srivastava, A. M., Doughty, D. A., and Beers, W. W. (1996). *J. Electrochem. Soc.* **143**, 4113.
- Srivastava, A. M., Doughty, D. A., and Beers, W. W. (1997). *J. Electrochem. Soc.* **144**, L190.
- Srivastava, A. M., and Sommerer, T. J. (1998). *Interface* **7**, 28.
- Srivastava, A. M., and Soules, T. S. (1995). "Kirk-Othmer Encyclopedia of Chemical Technology," Fourth Edition, Vol.15, p. 562, Wiley, New York.
- Stevens, A. L. N. (1976). *J. Lumin.* **12/13**, 97.
- Van Schaik W., et al. (1993). *J. Electrochem. Soc.* **140**, 216.
- Verstegen, J. M. P. J. (1974). *J. Electrochem. Soc.* **121**, 1623.
- Vij, D. R. (1988). "Luminescence of Solids," Plenum, New York.
- Weber, M. J. (1995). "Calorim. High Energy Phys., Proc. 5th Int. Conf." (H. A. Gordon and D. Rueger, eds.), World Scientific, Singapore.
- Wegh, R. T., et al. (1999). *Science* **283**, 663.
- Welker, T. (1991). *J. Lumin.* **48/49**, 49.



Photochromic Glasses

R. J. Araujo

Corning Glass Works

- I. Introduction
- II. Mechanism of Darkening and Fading
- III. Formation of Photochromic Glasses
- IV. Performance of Photochromic Glasses
- V. Benefits and Alleged Detriments

GLOSSARY

Chemical strengthening Strengthening of glasses by immersion in a bath of molten salt. An exchange of ions between the glass and the bath causes a thin layer of glass to be in compression and thus increases the physical strength of the glass.

Colloids Small aggregates of matter having properties differing slightly from those of the bulk material because of the increased importance of surface effects.

Color coordinates Numerical description of color determined by the ratios of transmittance at three different wavelengths.

CPF™lenses Ophthalmic lenses that because of their deep red color (strong blue absorption) provide a measure of comfort to some victims of *retinitis pigmentosa* and cataracts.

Dark adaptation Process by which vision improves with time after the eye has been subjected to a dramatic decrease in the level of lighting. This phenomenon depends on the regeneration of rhodopsin in the eye as well as on complicated neurological processes.

Photolysis Chemical or physical reaction induced by irradiation with light.

PHOTOCHROMISM is the phenomenon by which absorption of electromagnetic energy causes a reversible change in the color of a material. The word reversible implies that the system reverts to its initial state upon the cessation of irradiation. Certain glasses containing very small crystallites of silver halide exhibit this behavior.

I. INTRODUCTION

In 1964 the Corning Glass Works introduced to the marketplace a new family of ophthalmic lenses having the remarkable property of becoming dark in sunlight and reverting to their clear or colorless state indoors. Subsequently, both the Schott Glass Company in Germany and the Chance-Pilkington Glass Company in England began to manufacture these lenses. This family of glasses was invented by W. H. Armistead and S. D. Stookey, the latter being also the inventor of glass-ceramics, the material used in Pyroceram® nose cones and Corning Ware® cooking utensils. The phenomenon of changing color in response to sunlight irradiation is called photochromism, and materials showing such a property are said to be photochromic. Photochromism has been known since about

1880, and many hundreds of chemical compounds, most of them organic (that is, containing carbon) show this property. All the organic photochromic compounds share an unfortunate characteristic: They all fatigue, that is, they stop being photochromic after they have been darkened and faded a number of times.¹ The photochromic glasses discovered by Armistead and Stookey stand as a dramatic exception: They never fatigue. Small but significant changes in the speed of darkening and fading do occur during the first few cycles, but the degree of darkening and fading possible does not change even after a half a million darkening and fading cycles. This family of photochromic glasses owes its remarkable properties to a suspension of very tiny ($\sim 1\text{--}2$ millionth of an inch) crystallites of silver chloride. A similar family of photochromic glasses owes its properties to a suspension of tiny crystallites of cuprous chloride. The slight differences between the two families are discussed in Dotsenko *et al.* Because of the similarity between the two families, only the glasses containing silver halide will be discussed in this article. In the following sections, we discuss how these glasses are made and processed, the mechanism by which they darken and fade, and the manner in which conditions influence the performance characteristics of the glasses.

II. MECHANISM OF DARKENING AND FADING

To a very good degree of approximation, the darkening and fading of photochromic glasses can be ascribed to the photolysis of the tiny crystallites of silver halide that are suspended in a vast ocean of inert glass. Because of their minute size, the crystallites do not appreciably scatter light, so the glass appears to be quite transparent (i.e., not milky).

Figure 1 is a two-dimensional representation of the way in which positively charged silver ions (Ag^+) and negatively charged chloride (or bromide) ions (Cl^-) are arranged in a periodic fashion in a silver halide crystal in photochromic glasses. Note that singly charged copper ions occur in a few of the positive ion sites in **Fig. 1**. In a perfect crystal, there would be no such impurity, and silver ions would occupy all positive ion sites.

¹Organic materials developed recently by Pittsburgh Plate Glass fatigue only slightly. Spectacles containing these photochromic compounds can be used for several years with only slight loss of photochromic efficiency. They are sold under the trademark "Transitions." Organic materials with similar properties have been developed by Corning Incorporated and are incorporated in spectacles expected to be marketed soon.

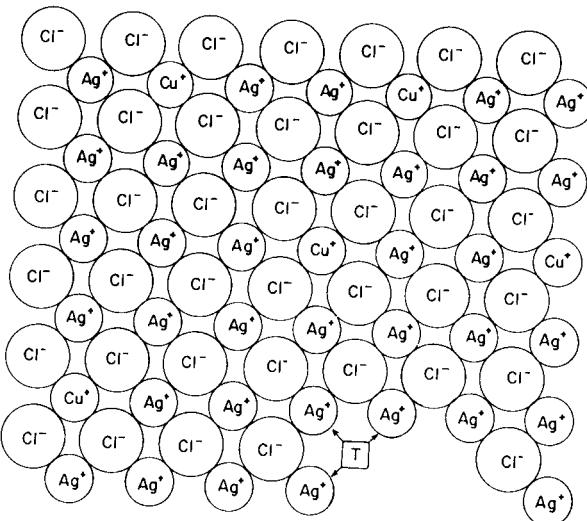
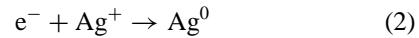
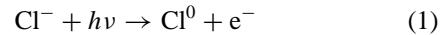


FIGURE 1 Silver halide before darkening.

If light of sufficiently high energy (ultraviolet (UV) or blue) impinges on the crystal, an electron can be transferred from a chloride ion to a silver ion:



Unless the resulting chlorine atom (Cl^0) and the silver atom (Ag^0) are moved apart extremely quickly, the electron simply returns to the chlorine atom, restoring the original state of the crystal, and the absorption of the photon has caused no net change. In spite of the rather crowded condition depicted in **Fig. 1**, a silver ion can move rather quickly from one positive-ion site to an empty neighboring positive-ion site even though, in doing so, it is squeezed terribly as it passes through the very small spaces between the chloride ions. Although the silver ion can move with surprising speed, it cannot move fast enough to prevent the electron from returning to its original chlorine atom parent; but fortunately, it does not have to. Another mechanism is available for separating the silver atom from the chlorine atom. (The chlorine atom does not move either, since chlorine atoms are much too big to move at room temperature.) Since all positive-ion sites in the interior of the crystal are equivalent, the energy of the system is not changed by an electron moving from one silver site to another. This motion of the electron effectively separates the silver atom from the chlorine atom. If the electron reaches a silver atom that is not equivalent to those on interior lattice sites (e.g., those marked **T** in **Fig. 1**), it becomes trapped there. This trapped electron (a Ag^0 on the surface) alternately attracts more electrons and silver ions until a speck of appreciable size forms, as can be seen in **Fig. 2**. This speck or colloid of silver

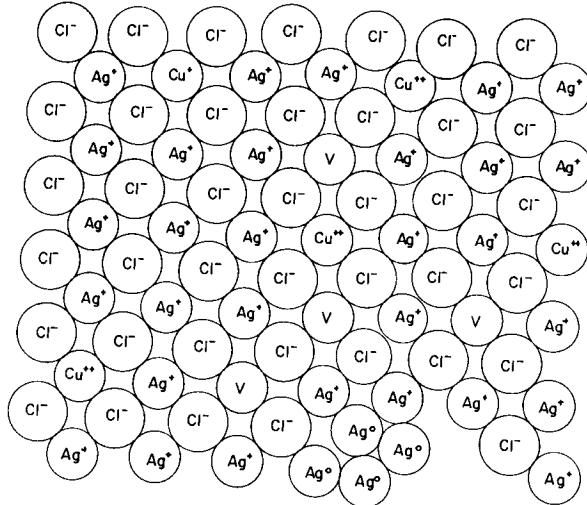
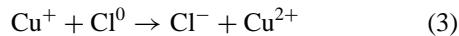


FIGURE 2 Silver halide after darkening.

is the entity that absorbs visible light in much the same way as does a metal and that causes the photochromic glass to appear dark. Note that the formation of the silver speck on the surface of the silver halide crystallite has caused some positive-ion sites to be vacant. Such sites are called vacancies and are designated by the symbol V in Fig. 2.

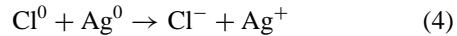
The process described so far is identical to that responsible for the formation of silver specks in photographic film. That is to say, in response to irradiation, electrons and silver ions form silver specks in both photographic film and photochromic glasses.

Note that no Cl^0 , formed according to Eq. (1), appears in Fig. 2. The reason is simple. Since all chloride ions in the interior of a silver halide crystallite occupy equivalent sites, the energy of the system is not changed when the Cl^0 borrows an electron from a neighboring chloride ion, effectively moving the Cl^0 from the original lattice site to a neighboring site. The site of a missing electron, that is, the site that has one electron less than it had in the state of lowest energy of the system, is called a hole. The hole is free to move from any lattice site to any equivalent neighboring site in the same manner as the electron. When the hole reaches a site next to a singly charged copper ion, it is at a site that is not equivalent to all the other negative-ion sites; so the hole, in effect, becomes trapped there. As a matter of fact, the singly charged copper ion gives an electron to its neighboring Cl^0 and becomes a doubly charged copper ion:

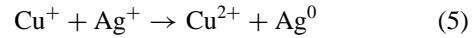


By this process the hole is said to be trapped by the copper ion. The doubly charged copper ion formed by the process is sometimes called a trapped hole.

It is fortunate that this process of hole trapping occurs. Otherwise, the hole would continue to wander in a random way until it came near the trapped electron (Ag^0), whereupon it would retrieve its electron, and the system would be restored to its initial state. Like the mobility (ease and speed of moving) of an electron, the mobility of a hole would be very high if it were not for the trapping; the process



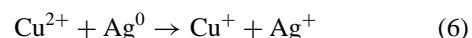
would then be so fast that colloid formation could not occur, and the photochromic glass would not darken. In effect, one vacancy and one doubly charged copper ion are formed for each silver ion that captures an electron and contributes to the growth of a silver speck. The net effect of Eqs. (1), (2), and (3) is given by



where Ag^0 is understood to be one of the atoms in the silver speck. Thus in Fig. 2, Cu^{2+} appears instead of Cl^0 . Vacancies occur because the silver ions that previously resided on those sites are now part of the colloid.

In photographic film, an analogous kind of hole trapping is required for colloid formation to occur. In this case, the Cl^0 chemically reacts with the gelatin in which the silver halide crystallite is suspended. Consequently, the hole is trapped in an infinitely deep trap. In no way can the process be reversed, and the silver speck or colloid is permanent. Thus, the darkening of a photographic film is said to be irreversible.

We have said that a photochromic glass returns to its initial clear state when irradiation ceases. That is, the darkening is said to be reversible. What accounts for the difference in behavior of a silver chloride crystallite in a photochromic glass and one in a photographic film is that in the glass there is no gelatin or any other substance with which the Cl^0 can react, other than with the singly charged copper ion. The doubly charged copper ion does not constitute an infinitely deep trap. The energetics of the situation strongly favor a process such as



which is obviously the inverse of darkening and is, indeed, the mechanism by which photochromic glass fades. The limitation on the speed of process (6) is the requirement that the doubly charged copper ion must be extremely close to the silver speck (one or two lattice sites) or the electron cannot make the jump. One can see from Figs. 1 and 2 that the doubly charged copper ions are not, in general, close enough for the electron to jump, that is, for process (6) to occur. How, then, does the fading take place?

The entire doubly charged copper ion is able to move from one lattice site to another (which, of course, must be vacant) in spite of the tight squeeze involved in getting past the big chloride ions. In fact, the motion of the doubly charged copper ion is much slower than that of the silver ion, precisely because of its double charge. This fact accounts for the observation that fading is much slower than darkening in photochromic glasses. Moreover, the speed of the copper motion depends more strongly, by far, on the temperature of the system than does the speed of the silver ion. The speed of electron motion or hole motion is virtually independent of temperature. Consequently, the speed with which the colloids form in response to radiation (speed of darkening of the glass) does not depend much on temperature; but the speed with which the glass fades depends very strongly on temperature. Of course, the glass fades more rapidly as the temperature is increased.

The statement that the electron cannot jump a long distance from the silver speck to a doubly charged copper ion must be modified somewhat. When the colloid is irradiated with visible light, there is a small but finite probability that a single electron will acquire so much energy that it can make a very long jump. This infrequent broad jump by the electron has the result that a photochromic glass fades somewhat faster when it is irradiated by long wavelengths (with no UV present) than it does in the dark. This slight acceleration of the fade rate by long wavelength irradiation is called optical bleaching. The fading that occurs even in the dark is called natural or thermal fading. The latter term is an attractive one because the rate of natural fading depends so strongly on temperature.

III. FORMATION OF PHOTOCHROMIC GLASSES

If salts of silver, copper, and chloride are added to the batch materials that are used to make glass, the resulting glass contains silver ions (Ag^+), chloride ions (Cl^-), and copper ions (Cu^+ or Cu^{2+}) in solution. Such a system is analogous to a solution of sodium chloride in water. When the silver, copper, and halide ions are completely in solution, the glass is not photochromic.

Fortunately, the amount of silver and chloride ions that can stay dissolved (solubility) depends very strongly on temperature for some glasses. At the high temperature at which the glass is melted, the silver and chloride ions are quite soluble; but in suitable host glasses the solubility decreases markedly with decreasing temperature. The fraction of the silver and chloride ions that are no longer soluble at low temperatures precipitates out in the form of silver halide crystals, and the material is then a photo-

chromic glass. Of course, if the host glass is such that the solubility of silver and chloride ions does not depend strongly enough on temperature, no silver halide precipitation occurs, and the glass is not photochromic, even though it contains silver and chloride. Thus not all glasses are suitable hosts.

Simply cooling the glass quickly to room temperature does not make it photochromatic. The glass must be heated to a moderately elevated temperature (but a low temperature compared with that of melting) for photochromism to be observed. At room temperature, even though the solubility of silver chloride is very low, the motion of the silver ions and chloride ions as they try to move through the glass looking for each other to form a silver halide crystal is exceedingly slow. At such speeds, the precipitation may take millions of years. At a moderately elevated temperature (but low enough so that the solubility of silver chloride is low) such as 600–700°C, the ions move fast enough to find each other and form droplets of molten silver chloride in about 20 min. (Silver chloride melts at about 455°C.) In most glasses, the mobility of chloride ions in the glass host is very slow at temperatures much lower than 500°C, and so growth of the silver halide droplet or crystal is virtually nonexistent at such low temperatures. Nevertheless, heating the glass in a temperature range below 500°C and much above room temperature can alter the photochromic properties. The reasons are not well understood, but some relevant observations can be made. Above 450°C the silver halide is molten. The molten droplet can be considered to be an infinitely defective crystal (the periodicity is destroyed). At temperatures slightly below the melting point, a large number of defects such as vacancies and kinks exist, such as those at position \square in Fig. 1. If the temperature is decreased slowly enough, many of these defects are annealed out. If the temperature is decreased too fast, many of the defects that are stable at high temperature persist at the low temperature, because the ionic motion required to correct the defects is too slow. Such phenomena can influence the performance of a photochromic glass. A second kind of phenomenon may also be influential on photochromic properties. Ions such as copper or lead ions can exist in a silver halide crystal, or they can exist in the glass. The exact manner in which they partition themselves between the silver halide droplets and the glass host depends on temperature. If the temperature is changed too rapidly, they may not have time to adjust to the partition in which they “want” to exist.

The reason the properties of a photochromic glass often change when the glass is chemically strengthened or tempered is probably related to phenomena similar to those just described. Ordinarily, nominal properties can be restored by repeating the heat treatment with more careful control of temperature and cooling rates.

IV. PERFORMANCE OF PHOTOCHROMIC GLASSES

For our purposes, we may think of light as being composed of small packets called photons. The energy of each of these packets is proportional to the frequency of the light, or equivalently, it is proportional to the reciprocal of the wavelength. Light of wavelength 350 nm is, therefore, more energetic than that of 550 nm wavelength. The reason green light does not cause typical photochromic glass to darken is that the energy of a photon of light with wavelength 550 nm is too low to remove an electron from a chloride ion and also give it enough energy to wander from silver site to silver site, as described in Section II. A photon of UV light, on the other hand, is quite energetic enough for this purpose.

When a photon of light is absorbed by the material through which the light is traveling, the intensity of the light is said to decrease. The intensity is conveniently thought of as the number of photons that pass through some unit area in a plane per unit of time. [Figure 3](#) represents a beam of light traveling from left to right and passing through a transparent but absorbing material M ; it has an intensity I_0 before passing through the material and a diminished intensity I after having passed through the material. The ratio of I to I_0 is called the transmittance. That is,

$$T = \frac{I}{I_0} \quad (7)$$

The transmittance of undarkened photochromic glasses is routinely measured using light of wavelength 365 nm. This transmittance is often expressed in percentage units and is called UV transmittance. Typical values range from around 30% in Photogray® lenses to as low as 5% in Photobrown® lenses. Another convenient number for expressing the fraction of the photons that pass through a

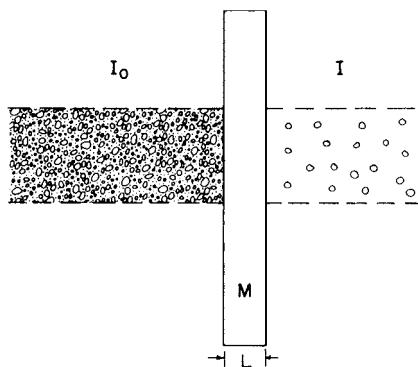


FIGURE 3 Representation of number of photons per unit area in a beam of light of intensity I_0 and the reduced number of photons in a beam of lower intensity I .

material is called the absorption coefficient, which is defined as

$$\alpha = \frac{1}{L} \ln \frac{I_0}{I} \quad (8)$$

where α is the absorption coefficient and L the thickness of the sample. A useful characteristic of the absorption coefficient is that it is proportional to the concentration of the species that is absorbing the light. The absorption coefficient of photochromic glasses rises very rapidly as the wavelength decreases below 350 nm. That is, photochromic glasses transmit very little UV. The transmittance in the visible portion of the spectrum of an undarkened photochromic glass is almost 90% unless colorants such as cobalt or nickel are added. When the glass is irradiated, the transmittance decreases with time and approaches an asymptotic or steady-state value. The speed with which the glass darkens during the initial period of irradiation is proportional to the intensity of the UV radiation. The proportionality constant depends on the UV transmittance. The rate of fading decreases considerably as the glass fades toward complete clarity. For our present purposes, we can describe the time dependence of the darkening process by

$$A = \frac{k_d I A_0}{k_d I + k_f} (1 - e^{-(k_d I + k_f)t}) \quad (9)$$

where A is the absorbance (αL), $k_d I A_0$ represents the speed of the darkening process (5), and k_f represents the speed with which process (6) occurs. When t is set equal to infinity in Eq. (9), the resulting expression describes the steady-state absorption,

$$A_{ss} = \frac{k_d I A_0}{k_d I + k_f} \quad (10)$$

As the intensity of the UV increases without limit, A_{ss} becomes equal to the limiting value A_0 . The way in which A_{ss} varies with I at constant temperature is the following: If I is sufficiently small so that $k_d I \ll k_f$, A_{ss} increases proportionately to I ; but if $k_d I \gg k_f$, A_{ss} virtually does not change with I .

Under actual sunlight conditions, commercially available photochromic glasses are in the regime in which A_{ss} changes only slightly with the intensity of UV. Between the hours of noon and 4:00 p.m., the intensity of UV may change by more than a factor of 10, but the transmittance of the photochromic glass changes by less than a factor of 2.

Temperature plays an important role in both the rate of fading and the determination of A_{ss} , the steady-state absorbance. A_{ss} is determined by the balance of forces trying to darken the glass and the forces trying to fade it. The principle is illustrated by a simple analogy. Consider

a very deep bucket with small holes uniformly distributed over its walls. Consider what happens when one tries to fill the bucket with water from a garden hose. The level of the water rises until the rate at which the water leaks out the holes exactly equals the rate at which the hose delivers water to the bucket. Then it stays constant. If one uses a larger hose that delivers water at a faster rate, the steady-state level of the water will be higher. The rate at which water leaks out at steady state will also be larger, because it can leak out a larger number of holes since the water level is higher. As soon as the hose is turned off, the level begins to fall because the leaking is no longer compensated. As the water continues to leak and the water level falls, the rate at which the level falls slows down, because fewer and fewer holes are involved in the leaking.

In photochromic glass, A_{ss} is analogous to the water level, the hose is analogous to $k_d I$, and the holes correspond to k_f , the fade-rate constant. If one raises the temperature, one increases the fade rate. In our analogy one can make the holes larger. This makes leaking faster and yields a lower steady-state water level for a given hose. In photochromic glass the speed of process (5) is virtually independent of temperature (like the hose). Process (6) depends strongly on temperature (as the leak rate depends on the size and number of holes); so A_{ss} , which depends on the balance, decreases with increasing temperature (just as the water level decreases with increasing hole size). That is, the higher the temperature, the less dark a photochromic glass can get. This is shown graphically in Fig. 4.

Although Eq. (9) and (10) are useful in describing the darkening of photochromic glasses in an approximate way, they fail to indicate some details. The fading rate in photochromic glass is not determined uniquely by how dark the glass becomes before irradiation was terminated (Fig. 5). If in two experiments performed at the same temperature, a glass is darkened (not to steady state) to the same value by UV irradiation of two different intensities, a faster fade rate is observed in the experiment in which the higher intensity is used.

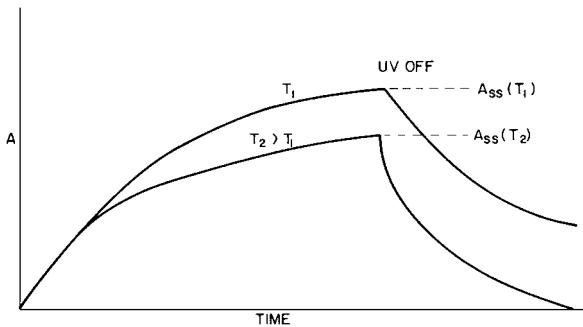


FIGURE 4 Darkening and fading as a function of time at two different temperatures.

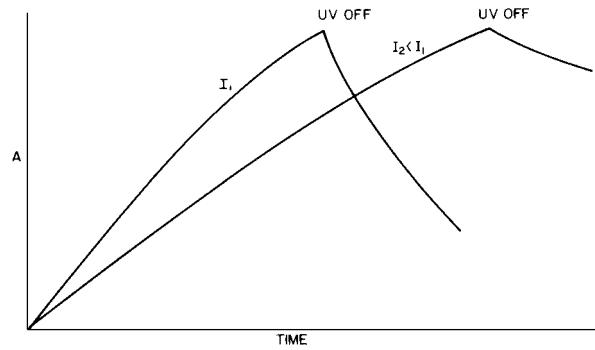


FIGURE 5 Short-time darkening at two different intensities and the corresponding fading.

Of course, had the glasses been darkened to steady state, somewhat more darkening would be observed when the higher intensity is used. In that case, the fade rate may have been faster or slower. Figure 6 illustrates that as one increases A_{ss} by increasing I over a certain range, the fade rate increases. If one increases A_{ss} by increasing I beyond that range, the fade rate begins to decrease again.

Often the glass is darkened at one temperature, and the temperature is changed when the irradiation is terminated (and when the glass can fade). For example, consider glasses that are darkened while the wearer is skiing on a very cold day. Only when the skier goes indoors does sunlight irradiation of the glasses cease, and only then do the glasses fade. It is quite likely that the temperature indoors is higher than that out of doors. The magnitude of the effect of changing temperature during fading varies from glass to glass but, in general, a glass darkened at a low temperature T_L and faded at a high temperature T_H fades faster than one darkened and faded at the same temperature, T_H . A glass darkened at T_H and faded at T_L fades more slowly than one darkened and faded at T_L .

The latter case is a common one. One wears sunglasses on a hot day. When one comes indoors, it is into an air-conditioned room. Because the situation is so common,

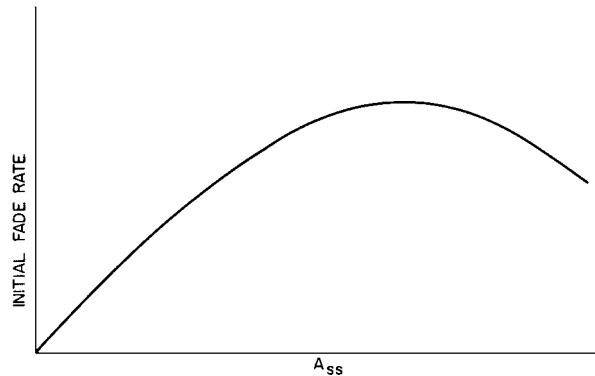


FIGURE 6 Initial fade rate as a function of the steady-state darkening level at constant temperature, A_{ss} , is varied by varying I .

Corning chooses to publish fade rates measured this way. Such published rates are the lowest likely to be observed. Other companies choose not to do this; the fade rates they publish would be higher than those published by Corning even if the glasses were identical.

It is worth noting that long-wave irradiation (red light) accelerates the fading process in many photochromic glasses. This fact is not of great importance when the glass is used for ophthalmic applications. However, it is relevant to other potential applications. For example, in some cases bleaching with polarized light causes the glass to become polarized. The mechanism by which this occurs is described by Araujo and Borrelli. Irradiation with high-intensity red light from a laser can lead to photodevelopment of a latent image produced by brief exposure of the glass to UV irradiation. This phenomenon is discussed by Dotsenko *et al.*

Although the manner in which the absorption spectrum of a silver colloid changes with size and shape is reasonably well understood, the factors that determine the size and shape assumed by the colloid when a photochromic glass darkens are not understood at all. Nevertheless, some general observations about color as a performance characteristic are in order. The darkened color of most photochromic glasses changes after the first few cycles. Photogray Extra® lenses, for example, darken to a brownish gray the first time they are exposed to light. After several darkening cycles (or a long period of continuous darkening), the color changes to the familiar blue-gray. An experimental glass known within Corning Glass Works as 15-15 darkens to a blue-gray when first exposed to UV. After several darkening cycles it darkens to a very bright red, similar to the color of Corning's CPF™ 550 lenses. A plot of the color coordinates measured at successive periods of time during the fading process may yield almost a straight line from the original color to the T_0 (original undarkened transmittance) color, but it may also yield a wild excursion in color-coordinate space. The darkened color and the color path during fading may both depend quite strongly on temperature. Photobrown™ lenses darkened to a brown color at temperature above 75°F, but at temperature below 32°F they were almost indistinguishable from Photogray® lenses.

Thickness also plays a key role in determining the color to which a photochromic glass darkens. In Photobrown® lenses, since the brown color depends so much on the UV transmittance of the glass, thick lenses yield a far better brown color than do thin lenses.

V. BENEFITS AND ALLEGED DETRIMENTS

Many tens of millions of pairs of sunglasses are sold in the United States every year, attesting to the desire of a large

number of people for protection from very bright light. For wearers of prescription lenses, the advantages of a single pair of glasses that function as sunglasses outdoors and clear glasses indoors from the simple point of view of convenience is self-evident.

There is no disagreement among vision researchers that UV and blue light can cause damage to both the anterior eye and the retina when the intensity reaches the level equivalent to looking directly at the sun at its apex. There is some opinion that blue light is somewhat hazardous at considerably lower levels, but most researchers believe that the danger decreases sharply as the intensity decreases.

Because of the natural photophobia of the normal person, it is unusual for the eye to be exposed to those levels of blue light or UV that are hazardous. Exceptions can occur, however. Reflections from snow fields or white sands can produce, at the cornea, intensities up to one-third the intensity produced by looking at the sun. Specular reflections (those observed on a mirror) from chrome ornaments or glass surfaces with reflecting coatings can lead to enormously high intensities. It is, of course, impossible to know the danger such random circumstances present to the average person. The "prudent man" concept, however, would suggest that glasses that absorb a substantial fraction (90%) of potentially harmful rays are a benefit.

There is a substantial body of thought that total deprivation of blue and UV constitutes a health hazard. While no general consensus exists on this point, it seems unwise as well as unnecessary to strive to remove 100% of the blue light from reaching the eye. Although further research may change expectations, at the present moment it is believed that a dark gray glass (between 1 and 30% luminous transmittance) that absorbs more than 90% of all the radiation below 380 nm is a benefit. Although photochromic glasses normally do not darken below 10% transmittance, they are, nevertheless, remarkably well suited for eye protection as it is currently understood.

It is important to note that in this context Photobrown Extra® lenses are to be regarded as gray. The most important advantage of the photochromic glasses over other glasses, which conceivably can be designed to have the same absorption spectrum, is that the convenience afforded by the automatically adjusting transmittance makes them more likely to be worn than glasses that must be exchanged when the environment changes.

A bright environment slows down the rate at which the eye can dark adapt. This effect is cumulative. A lifeguard, for example, who spends 8 hr on the beach every day for a week does not regain his normal nighttime vision even after 24 hr in the dark. A person who wears dark glasses in the sun dark adapts to low-level light more quickly than one wearing no glasses. If, in addition, the dark glasses are photochromic and begin to clear up when the light

intensity is significantly reduced, an added, albeit small, benefit exists.

Because photochromic glasses darken most strongly in cold weather and because the fading rate observed at room temperature is maximal for a glass darkened at very low temperatures, the advantages of photochromic glasses are certainly maximized for people like skiers who spend time outdoors in cold weather.

The question of color distortion of Photobrown® lenses has been occasionally raised. The absorption coefficient induced in this glass during darkening is really quite flat (the human eye is a remarkable detector of small differences); so although very slight color changes are perceivable when comparisons are made, absolutely no danger of color misidentification exists for the normal eye.

It has been suggested that the automatic response of photochromic glasses renders a changing pupil size unnecessary and, therefore, leads to atrophy of the iris. Virtually all the variations in size of the pupil at steady state occur at light levels considerably lower than those found in the typical office or classroom. Since no sunglasses of any kind are normally worn under such conditions, they cannot have an effect. Furthermore, the speed of response of photochromic glasses is much slower than the speed of pupillary response. Therefore, even if photochromic glasses were worn and even if they did respond to the same range of intensities to which the pupil responds, they would not replace pupillary response because they are too slow.

Sunglasses yield comfort and perhaps a measure of eye protection. The eye protection is probably improved if the sunglass absorbs UV and short-wave blue rather strongly. Photochromic glasses do this very nicely, and

they also provide the convenience of automatic adjustment. The wearer of photochromic glasses is not likely to forget to don his eye protectors when venturing into the outdoors.

SEE ALSO THE FOLLOWING ARTICLES

COLOR SCIENCE • GLASS • PHOTOCHEMISTRY, MOLECULAR • PHOTOGRAPHIC PROCESSES AND MATERIALS

BIBLIOGRAPHY

- Araujo, R. J. (1968). *Appl. Opt.* **7**, 781.
- Araujo, R. J. (1971). In "Photochromism" (G. H. Brown, ed.), p. 680, Wiley, New York.
- Araujo, R. J. (1977). In "Treatise on Materials Science," Vol. 12. (M. Tomozawa and R. H. Doremus, eds.), p. 91, Academic Press, New York.
- Araujo, R. J. (1980). *Contemp. Phys.* **21**, 77.
- Araujo, R. J. (1980). *J. Non-Cryst. Solids* **42**, 209.
- Araujo, R. J. (1982). *J. Non-Cryst. Solids* **47**, 69.
- Araujo, R. J., and Borrelli, N. F. (1991). In "Optical Properties of Glass," (Kreidl and Uhlmann, eds.), p. 125, American Ceramic Society, Westerville, Ohio.
- Araujo, R. J., Borrelli, N. F., and Nolan, D. A. (1979). *Phil. Mag.* **B40**, 279.
- Araujo, R. J., Borrelli, N. F., and Nolan, D. A. (1981). *Phil. Mag.* **B44**, 453.
- Dotzenko, A. V., Glebov, L. B., and Tsekhomsky, V. A. (1997). "Physics and Chemistry of Photochromic Glasses," CRC Press, Boca Raton, FL.
- Hoffman, H. J. (1990). In "Photochromism" (Durr and Bouas-Laurent, eds.), p. 822. Elsevier, Amsterdam.
- Nolan, D. A., Borrelli, N. F., and Schreurs, J. W. H. (1980). *J. Am. Ceram. Soc.* **63**, 305.



Sandwich Composites

Charles E. S. Ueng

Georgia Institute of Technology

- I. Brief History
- II. Concepts and Functions
- III. Classifications
- IV. Applications
- V. Manufacturing Aspects
- VI. Mechanical Aspects
- VII. Problems in the Use of Sandwich Composites
- VIII. Recent Developments
- IX. Future Trends

GLOSSARY

- Anisotropy** Direction-dependent mechanical property.
Anticlastic Saddle-shaped.
Core Middle layer, relatively soft, light weight, and thick.
Facings Outer layers, relatively strong, heavy, and thin.
Isotropy Direction-independent mechanical property.
Orthotropy Material symmetry property, with respect to two or three mutually perpendicular planes.
Superplastic Material characteristic, able to have a very large stretching.

SANDWICH COMPOSITES are made from two thin, strong layers of facing sheets and one thick but weak layer of core filled between the two facing plates. Under perfectly bonded conditions, they form a structural component that is capable of providing a high strength-to-weight ratio and a high stiffness-to-weight ratio for the fulfillment of lightweight structural design purposes, such as

for aerospace vehicles and ground transportation, packaging, and other engineering applications. With the advent of fiber-reinforced composite during the past three decades, a sandwich composite with facings made of thin laminates can further increase the effectiveness of modern advance composites through the concept of sandwiching them together.

I. BRIEF HISTORY

People have been looking for light but strong materials ever since they began to use devices for transportation. After the Wright brothers invented the airplane, the search of such materials intensified because it is extremely important that light and strong materials be used for the design of a wing and fuselage structure in order for an airplane to provide a high payload under a given thrust force from the propelling system.

When aircraft were first being built, wood was used extensively. It formed a monocoque, or unstiffened fuselage,

in which three layers of wood having different orientations were glued together. By the end of the 1920s, lightweight aluminum alloys became the major material for transport aircraft construction. The speed of the airplane gradually increased during the 1930s, and the smoothness of the airflow over the surface became even more important since any buckling of the surface causes an inadmissible rise in drag at high speeds. New types of construction that do not buckle under load had to be developed.

During World War II, the famous British Mosquito airplane was manufactured by adopting the sandwich idea with the use of lightweight balsa wood. Since it increases stiffness tremendously without adding too much weight, sandwich construction has been used ever since, not only in aircraft and space vehicle design, but also for ground transportation, building construction, packaging, and other engineering applications.

II. CONCEPTS AND FUNCTIONS

A sandwich structure consists of two face plates separated by, and attached to, a core, as shown in Fig. 1. In general, the facings (or face sheets) are much thinner and possess greater stiffness and strength than the low-density core. The facings are usually made of thin, stiff, strong sheets of solid material such as metal, alloy, plywood, reinforced plastic, or cardboard. During recent years, carbon filament reinforced epoxy resin panels have become commonly used for face plates, particularly in the aerospace industry. The core materials, on the other hand, can vary very widely from those used for the facings. Cores are often made from metallic, plastic, fibrous, or felted materials, and have a multitude of configurations ranging from corrugated to cellular to foamed structures. The concept behind sandwiched construction is to have the facings and core act in concert as a very efficient structural element.

The function of a sandwich composite is analogous to an I beam, which has been used widely in structures for many decades. It is well known that the purpose of the web of an I beam is to bear the shear force, while the two flanges, separated by the web, provide the bending rigidity. The core in a sandwich composite serves the same function as the web of an I beam in that greater flexural rigidity is obtained from the facings. In order to accomplish this, the

core must be stiff enough to ensure that the facings remain at a proper distance apart. The core must also provide adequate shearing strength so that the facings will not slide relative to each other when the sandwich panel is bent. In the absence of necessary shear strength, the two thin facings would act as two independent beams or panels, and lose the sandwich effect. Finally, the core must also possess enough stiffness so that the facings will stay flat or nearly flat when they are subjected to compressive stresses that would otherwise cause buckling or wrinkling.

The objective of the sandwich composite is to offer a structure that is strong and stiff but at the same time lightweight. This special feature plays a very important role in the design of aerospace vehicles, such as airplanes, rockets, and spaceships. The major advantages of sandwich composites over conventional materials are that sandwich composites (1) have a low overall density, a high strength-to-weight ratio, and a high stiffness-to-weight ratio; (2) are capable of providing good thermal and acoustical insulation; and (3) have uniform energy absorption capacity. Such overall versatility has contributed greatly to the development of lightweight sandwich composites.

III. CLASSIFICATIONS

Having defined what a sandwich composite is, one can therefore classify such a structure according to the type of core material and the type of facing material used. For example, Fig. 2(a) shows a sandwich composite consisting of foam-type core, Fig. 2(b) represents a honeycomb core, Fig. 2(c) involves a unidirectional corrugated core, and Fig. 2(d) shows a variation of the unidirectional corrugated core, with back-to-back corrugation. Here the corrugated core includes two fluted metal sheets attached. The honeycomb-type core is usually manufactured from aluminum foil. The ordinary honeycomb core tends to bend anticlastically and will not readily fit a cylindrical surface. An improved version, the flexible core with multiwave, seems to provide exceptional formability into compound curvatures. It also provides higher shear strengths than does a comparable hexagonal core of equivalent density. A superplastic forming technique has been applied for manufacturing new configurations of core shapes. Preliminary results with this new method seem to be very promising.

Facing sheets can be made from aluminum, stainless steel, or other alloys. Fiber-reinforced laminates, such as glass-epoxy, graphite-epoxy, and boron-epoxy, have been used frequently as facing sheet materials. On the other hand, facings made of plywood, glass-reinforced cement, plasterboard, resin-impregnated paper, hardboard, and other materials are frequently used in building construction.

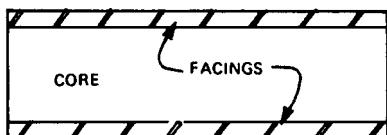


FIGURE 1 Basic sandwich structure.

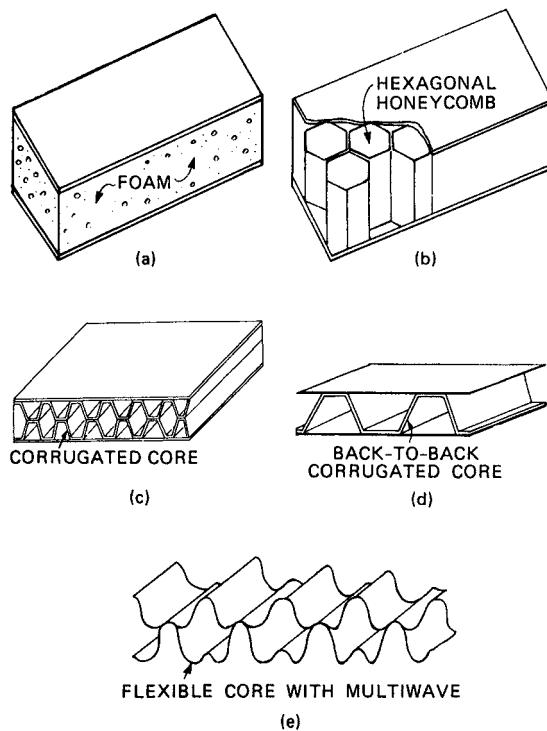


FIGURE 2 Types of core materials in sandwich composites: (a) foam, (b) honeycomb, (c) corrugated, (d) back-to-back corrugated, and (e) flexible.

Depending on the application of a particular sandwich composite, almost any type of core can be glued together with any type of facing sheet. Sandwich composites can be manufactured for specific functions, but, of course, commercially available composites cost much less than custom-made materials.

IV. APPLICATIONS

Because of their high strength-to-weight and high stiffness-to-weight ratios, modern sandwich composites can be applied to many engineering problems wherever such advantages are necessary or desirable. The following is a list of such applications for which sandwich composites have been commonly employed. Note that this list is by no means exhaustive and their new applications come up periodically.

1. Aerospace structures: Sandwich composites are used almost everywhere in aerospace structures, from small private airplanes to huge transport C-5Bs and from rockets to the space shuttle. Wing structures, fuselages, rotor blades, rocket cases, heat shields, floors, and passenger compartments are just a few examples. A single C-5B military

transport contains more than one acre of sandwich composite construction, which supports a substantial advantage in the payload.

2. Building construction: Lightweight sandwich composites can provide an advantage over conventional construction materials. When acoustic insulation is needed, a sandwich partition or a door made of sandwich construction can perform the function very well without the extra depth and weight usually required. This can often leave more usable space in a building. The reduction of weight could be rather significant for a high-rise structure. Such a weight reduction may also lead to a smaller column size and finally a lighter foundation. In a tall building located in an active seismic zone, a reduction of structural weight will immediately contribute to a decrease of seismic forces involved. For long-span structures, sandwich composites may open up a new dimension in design for architects and engineers.

3. Ground transportation vehicles: In order to improve fuel efficiency, the weight of ground transportation vehicles must be kept to a minimum. Sandwich composites can serve this purpose. In addition to the advantages sandwich construction offers such as high strength-to-weight and high stiffness-to-weight ratios, it can also absorb a substantial amount of energy in traffic accidents. The ability to serve as an insulation layer to avoid unwanted heat transfer and as a noise barrier are attractive features particularly important for large transportation vehicles such as rail cars and buses.

4. Packaging industry: Sandwich composites, primarily those made of paper, are widely used in the packaging industry. As long as they are kept dry, packaging boxes made of composites material can serve very well. Specially treated paperboard can provide a minimum water absorption rate to meet certain military requirements. It is one of the most widely and universally adopted materials in the packaging industry.

5. Freight containers: For low-cost land and sea transportation purposes, the use of freight containers has become a fairly standard approach. During recent years, the same idea has also been extended to air cargo industry after the wide-body cargo airplanes came into use in the freight transportation industry. Because of the reusable nature of these containers, they must be rugged and structurally strong, and at the same time, they should be lightweight. Sandwich composites with fiber-reinforced plastics as facing sheets and rigid structural foam or plywood as core are preferred for the construction of such containers.

6. Ocean engineering industry: The materials most commonly used as cores for sandwich panels for marine use include foam, end-grain balsa, plywood, and honeycomb. Any of the core materials listed here with glass-reinforced plastic skins can be used. Sandwich

construction is used primarily for large, relatively flat, unstiffened surfaces such as decks, cabin tops, and bulkheads, although a number of sandwich hulls have been built.

V. MANUFACTURING ASPECTS

As mentioned earlier, a sandwich composite is made up with one thick layer of core in the middle and two thin layers of facing sheets bonded together. The manufacturing process for each of these can be briefly stated as follows.

1. Core: the most popular type of core used today is in the hexagonal shape, also known as the honeycomb. Honeycomb is primarily by the expansion method shown in a diagram in Fig. 3. It begins with the stacking sheets of aluminum or other materials on which the adhesive nodal lines have been printed. The next step is to cut the block into the required dimension—that is, the thickness of the core. A honeycomb core is obtained by expanding it in the vertical direction shown. For higher-density cores, the sheets are corrugated first, and then the adhesive is added to the corrugated nodes and the sheets are stacked into blocks. Finally, the core panels are cut from these blocks to the required thickness.

2. Facing sheets: Metallic facing sheets are simply cut from commercially available products, but for nonmetallic facing sheets, the manufacturing process is much more complicated. Usually, it starts with “prepregs”—a form of preimpregnated fibers in which the fibers run in the lengthwise direction of the tape. According to the desired orientations obtained from analysis, a lay-up is followed for a particular stacking sequence. Then a drying or polymerization process, known as curing, is performed in an autoclave so that the resinous matrix material forms a permanent bond between the fibers and between the individual fibers. The curing process must be done under proper temperature, pressure, and time control in order to ensure that the quality of laminates meets industry’s rigid requirements.

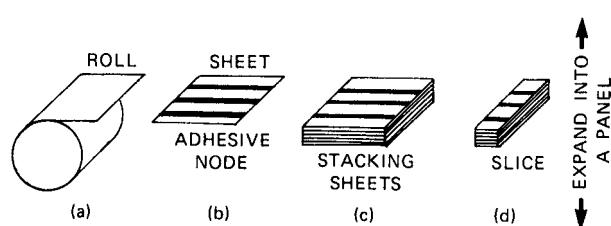


FIGURE 3 Expansion method for manufacturing honeycomb cores.

With the help of advanced stress analysis results and the sophisticated software development that have become available during the last few years, today’s computer-controlled manufacturing machines can place the fibers along the required orientation at a special location for a particular structural segment. Automation cuts down the manufacturing labor cost by a significant proportion and makes sandwich composites more competitive than ever before.

VI. MECHANICAL ASPECTS

Due to the substantially different material properties of the facings and the core, a sandwich composite is obviously a heterogeneous and possibly also anisotropic in nature. The complexity and difficulty involved in mechanical analysis of such a structural element are indeed considerably more involved than the homogeneous and isotropic case as carried out in the classical analysis. That the core material itself is not a solid continuum but a hollow one, immediately causes a mathematical analysis to be more challenging. In order to make the analysis manageable, certain simplifications must be made. The following assumptions are frequently adopted.

1. The facings are very thin in comparison with the total thickness of the sandwich, and consequently they are treated as elastic, isotropic, or orthotropic thin plates (i.e., plane sections initially perpendicular to the middle plane of the plate remain planar after deformation in accordance to the Bernoulli–Navier hypothesis). Only three stress components, σ_x , σ_y , and τ_{xy} are considered for the facings, where the $x-y$ plane represents the plane of the plate, and σ and τ indicate the normal and shear stresses, respectively.

2. The core consists of an elastic, orthotropic (in certain cases, isotropic) continuum whose load-carrying capacity in the plane of the sandwich is negligible.

3. The modulus of elasticity of the core in the direction perpendicular to the facings is infinite, but the shear modulus of the core in the plane of the sandwich can be ignored.

4. The core is assumed to be subjected to a state of antiplane stress that can be defined as the state of stress that exhibits stress components that are zero in a state of plane stress. That is, only σ_z , τ_{xz} , and τ_{yz} are involved in the core.

5. Perfect continuity exists at the interfaces at which the facings and core are bonded together.

In addition, there may be certain other assumptions made on an ad hoc basis for special cases.

The most important difference in the derivation of the governing differential equations between a sandwich composite and a homogeneous counterpart is that the shear strains in the core of a sandwich construction may not be ignored, since the core has a low modulus of rigidity in the plane of the sandwich. The resulting differential equations that characterize the mechanical behavior are always more complicated than the single-layer homogeneous cases. The difficulty in seeking a mathematical solution is obviously more involved. In many cases, only approximate solutions are possible.

VII. PROBLEMS IN THE USE OF SANDWICH COMPOSITES

The following problems are often encountered in the use of sandwich composites.

1. Joining problems: As with any composite material structure, the joining problem is a more difficult task than with conventional materials. Adopting either an adhesively bonded approach or a mechanically fastened approach requires special attention to the preparation and treatment of the components and understanding of the problem before a successful joining task can be accomplished. An effective and efficient joint should not be overdesigned since this could erode advantage of the weight saved from the use of sandwich composites.

2. Analysis problem: As mentioned before, the mathematical equations characterizing the sandwich composites are considerably more difficult than their counterparts with which only single layer is involved. Generally speaking, they contain a larger number of differential equations, of more variables and higher order, and often are coupled. Approximate solutions may be necessary for most of the practical problems in this category.

VIII. RECENT DEVELOPMENTS

The following new developments on applications of sandwich construction have become available during recent years.

1. Aeronautical applications: In recent aviation history, first, two pilots flew the *Voyager* airplane nonstop around the world in December 1987 without refueling. It opened up a new page of technological history in aeronautical engineering. Second, in July 1989, the U.S. Air Force revealed the real but previously unconfirmed Stealth Bomber which is able to evade enemy's radar detection. The two craft belong to two completely different categories of airplane design in terms of size, function, speed, and capabil-

ities, but according to the open literature and news media, the concept of composite sandwich was adopted in cooperation with laminate technology in the structural designs. Once again, those two events demonstrated that many advantages of sandwich construction can be used to fulfill the special requirements of such projects.

During the last few years, such advancements in sandwich composites have entered into the business jets manufacturing industry. The capabilities of such small jets fall some where between the last two extremes, but able to seat six to twelve people and to travel across the ocean without refueling at the speed comparable to a commercial big jet. Such airplanes can bring corporation executives from point to point around the globe in a very time-saving fashion. Without the application of lightweight sandwich composite technology, it would not be possible to have made this aviation dream to become a reality.

2. Marine applications: It was reported that in the 1988 America's Cup event, the hull and the deck of both *Star and Stripes* and *New Zealand* were primarily of laminated sandwich construction. Although the latter is twice as long as the former, 39.7 versus 18.3 m, with a weight ratio of nearly 10 to 1. The yachts were designed by two different teams. Due to the competitive situation, the detailed processes were not known to other side during the time the yachts were made. The characteristics on porosity, stiffness, impact strength, and manufacturing methods were the main considerations.

The fiber-reinforced plastics (FRP) sandwich construction has also been widely used in small- to medium-sized boat designs, military as well as civilian, such as minesweepers and ferry boats, particularly popular in Australia and the Scandinavian countries.

The concept of using very thick sandwich composites for submarine hull construction has been reported in the open literature during recent years. Obviously, they can provide certain advantages over conventional shipbuilding materials.

3. Construction and infrastructure applications: Construction is an area where momentum has recently picked up because the FRP unit cost has dropped, and it shows such advantages over metals as corrosion resistance and freedom from magnetic effects. Many large-scale applications are being expanded into building construction, highway bridges, and off-shore structures. They include not only new constructions but also many rehabilitation projects.

Another encouraging indicator for the popularity of sandwich structure is the founding of a technical journal exclusively devoted to the study of sandwich constructions, known as *Journal of Sandwich Structures and Materials* and published by Technomic Publishing

Company from 1999. It attracts research workers around the world to report their research findings in this important area. In addition, there have been four international conferences on sandwich construction, with the first one held in 1989 and three more after that spaced at every three-year interval. From the contents of this journal and the conference proceedings, one can easily be convinced that sandwich composites will not only continue to play its important role but also to accelerate its applications in the twenty-first century's world.

IX. FUTURE TRENDS

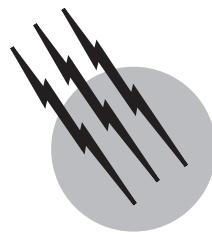
Because of the superior mechanical properties of sandwich composites mentioned in this article, it is anticipated that a wider application of these versatile structural components is in sight of the future. With the help of high speed and huge memory computational facilities available today, engineers can now take a smaller and smaller portion of such a complicated structural component, particularly the core part, and carry out a more refined mechanical property analysis. In addition, advanced computerized manufacturing process can also reduce the production cost and make it more competitive for more engineering applications.

SEE ALSO THE FOLLOWING ARTICLES

ADHESION AND ADHESIVES • CARBON FIBERS • COMPOSITE MATERIALS • CONCRETE, REINFORCED • MASONRY • MECHANICS OF STRUCTURES • METAL MATRIX COMPOSITES • SPACECRAFT STRUCTURES

BIBLIOGRAPHY

- Allen, H. G. (1969). "Analysis and Design of Structural Sandwich Panels," Pergamon, New York.
Dietz, A. G. H., ed. (1970). "Composite Engineering Laminates," MIT Press, Cambridge, MA.
Norton, B. B., ed. (1974). "Composite Materials: Engineering Applications of Composites," Vol. 3, Academic Press, New York.
Plantema, F. J. (1966). "Sandwich Construction," Wiley, New York.
Ueng, C. E. S. (1978). "Superplastically formed new sandwich cores," *Trans. Eng J. TE4*, 437–447.
Ueng, C. E. S., and Kim, T. D. (1983). "Shear modulus of core materials with arbitrary polygonal shape," *Int. J. Computers Structures* **16**, 21–25.
Ueng, C. E. S., Underwood, E. E., and Liu, T. L. (1979). "Shear modulus of superplastically formed sandwich cores," *Int. J. Computers Structures* **10**, 393–397.
Vinson, J. R. (1999). "The Behavior of Sandwich Structures of Isotropic and Composite Materials," Technomic, Lancaster, PA.



Silicone (Siloxane) Surfactants

Randal M. Hill

Dow Corning Corp.

- I. Introduction
- II. Molecular Structures
- III. Synthesis and Chemistry
- IV. Surfactancy
- V. Applications

GLOSSARY

Emulsion A dispersion of two immiscible liquid phases such as oil and water. Requires a surfactant for stability.

Liquid crystal (lyotropic) Self-assembled microstructured phase formed by surfactants. Long-range ordering is present in combination with properties characteristic of liquids such as fluidity.

Micelle Globular self-assembled aggregate of surfactant molecules in a solvent.

Microemulsion Spontaneously self-assembling transparent mixture of two immiscible phases and a surfactant.

Polyether Polyoxyethylene, polyoxypropylene, polyoxybutylene, or copolymers thereof.

Silicone An organofunctional siloxane. Examples include polydimethylsiloxane and polyphenylmethylsiloxane.

Surfactancy The collection of properties characterizing a surfactant—includes surface and interfacial tensions, self-association and liquid crystal phase behavior, wetting, foaming, emulsification, and colloid stabilization.

Surfactant A surface active substance containing spatially separated polar and nonpolar groups.

Vesicle Closed capsule structure consisting of one or more sheets of surfactant bilayer.

Wetting Spreading of one phase over another. May be spontaneous or forced.

SILICONE (OR SILOXANE) surfactants are surface active agents consisting of a permethylated siloxane group coupled with one or more polar groups. The most common examples are graft copolymers in which several poly(oxyalkylene) groups are grafted onto a polydimethylsiloxane backbone. This class of surfactants is used in a wide range of applications in which conventional surfactants are ineffective. They are surface active in aqueous and nonaqueous media. They self-associate in solution to form micelles, vesicles, liquid crystal phases, and microemulsions (with oils). Their uses include plastic foam stabilization, as wetting and spreading agents and lubricants, as water-in-oil emulsifiers, and to impart a silicone feel to personal care products.

I. INTRODUCTION

Silicone surfactants consist of a permethylated siloxane group coupled with one or more polar groups. Three of their unusual properties are

- a. They are surface active in both water and in nonaqueous media such as mineral oil and polyols.
- b. They are able to lower surface tension as low as 20 mN/m, some 10 mN/m lower than is achievable using conventional surfactants.
- c. They remain liquid to very high molecular weights.

All of these properties may be traced to the characteristics of the silicone block—the unusually high flexibility of the siloxane backbone (high T_g, liquid at high MW), and the presence of methyl ($-\text{CH}_3$) groups giving very low surface energy. The $-\text{O}-\text{Si}-\text{O}-\text{Si}-$ backbone appears to serve mostly as a flexible framework for the methyl groups. The surface energy of a methyl-saturated surface is about 20 mN/m, and this is also the lowest surface tension achievable using a silicone surfactant. In contrast, most hydrocarbon surfactants contain alkyl, or alkylaryl hydrophobic groups that are mostly $-\text{CH}_2-$ groups, and pack loosely at the air/liquid interface. The surface energy of such a surface is dominated by the methylene groups and for this reason hydrocarbon surfactants typically give surface tensions of about 30 mN/m or higher.

The surfactancy properties of polymeric silicone surfactants are notably different from those of common hydrocarbon polymeric surfactants, such as the EO/PO block copolymers. Comparable silicone surfactants usually give lower surface tensions, and their aggregation properties in aqueous solution are strikingly different—silicone surfactants often form bilayer phases and vesicles rather than micelles and gel phases. The skin feel and lubricity properties of silicone surfactants do not appear to have any parallel among hydrocarbon polymeric surfactants.

Although there are important differences with conventional hydrocarbon surfactants, silicone surfactants share many features in common with these better-known surfactants. Equilibrium and dynamic surface tension vary with concentration and molecular architecture in similar ways. Silicone surfactants self-associate in solution to form micelles, vesicles, and liquid crystal phases. Self-association follows similar patterns as molecular size and shape are varied. Silicone surfactants containing polyoxyalkylene groups exhibit an inverse temperature solubility or cloud point. HLB values can be calculated for silicone surfactants, but more useful values can be obtained from calculations that take into account the differences between silicone and hydrocarbon species.

It is often stated that silicones and silicone surfactants are lipophobic—that is, they are incompatible with hydrocarbon oils, but this is only partly correct. Small silicone surfactants, such as the trisiloxanes, are very compatible with organic oils. For example, aqueous solutions of the

trisiloxane surfactants give low interfacial tensions against alkane oils. The incompatibility between polymeric silicone surfactants and some hydrocarbon oils is due to the polymeric nature of the silicone block¹ rather than to phobicity such as that between fluorocarbon and hydrocarbon groups.

Silicone surfactants were first introduced in the 1950s for the manufacture of polyurethane foam. Soon afterwards other applications were invented for them. Non-aqueous surface activity is the basis for their use in polyurethane foam manufacture and as demulsifiers in oil production and foam control agents in fuels. Ability to lower surface and interfacial tension leads to wetting and spreading applications. Unusual molecular structures and high molecular weights make them useful as novel emulsifiers. Silicones impart a unique dry-lubricity feel to surfaces such as textiles, hair, and skin. Since silicone surfactants incorporate silicone in a water-soluble or water-dispersible form, they represent a convenient means for delivering silicone lubricity or feel to a surface by way of an aqueous formulation.

Use of silicone surfactants continues to expand and find new applications. Substantial advances in our understanding of this class of surfactants have taken place in recent years including an understanding of their aqueous aggregation behavior, their ternary phase behavior with silicone oils, and their ability to promote rapid wetting of hydrophobic substrates. This overview will describe the structure, preparation, and surfactancy properties of this fascinating class of surfactants. A brief discussion of some common applications will also be given to illustrate how the unusual properties of silicone surfactants are used. Although a significant amount of information regarding uses of silicone surfactants is still found primarily in the patent art, the major applications have now been reviewed in journal articles.

Other names for them include silicone polyethers (SPEs), polyalkylene oxide silicone copolymers, silicone poly(oxyalkylene) copolymers, and silicone glycols. The International Cosmetic Ingredient Nomenclature and the CTFA adopted name is dimethicone copolyol.

II. MOLECULAR STRUCTURES

The hydrophobic portion of silicone surfactants is a permethylated siloxane group. Many different types of

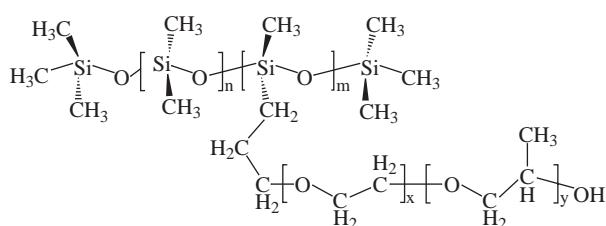
¹The mutual solubility of two molecular species decreases with increasing molecular weight. For example, the solubility of PDMS in hexane decreases with increasing MW of the PDMS. For the same reason, higher MW silicone surfactants are less compatible (but may be more surface active) with organic oils.

hydrophilic groups can be incorporated—some examples are given in the following table.

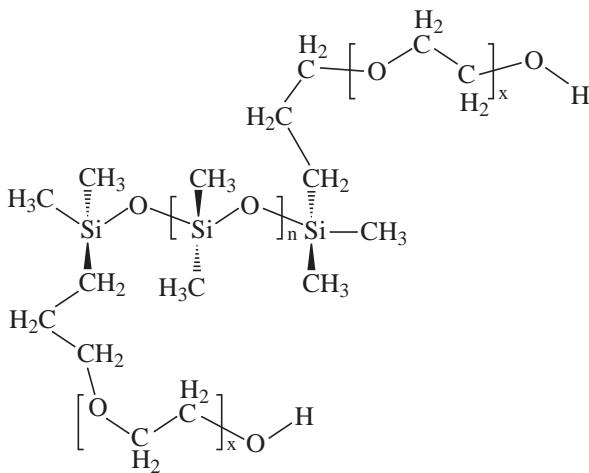
Types and Examples of Polar Groups in Silicone Surfactants

Type of polar group	Examples
Nonionic	Polyoxyethylene Polyoxyethylene/polyoxypropylene
Anionic	Carbohydrates
Cationic	Sulfate
Zwitterionic	Quaternary ammonium salts
	Betaines

Silicone surfactants may be either polymeric or non-polymeric types. The most common polymeric silicone surfactants are graft-type (rake-type) or ABA copolymers. Structures illustrating these types of copolymers are shown in the following diagrams.



Graft-type silicone polyoxyalkylene copolymer containing both EO and PO groups, $M'D'E_xP_yOH)_mM$.

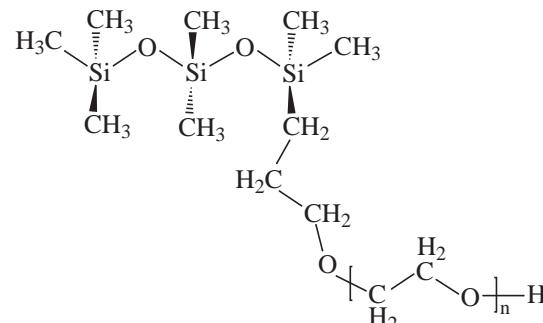


ABA-type silicone polyoxyethylene copolymer, $HOE_xD_{n+2}E_xOH$.

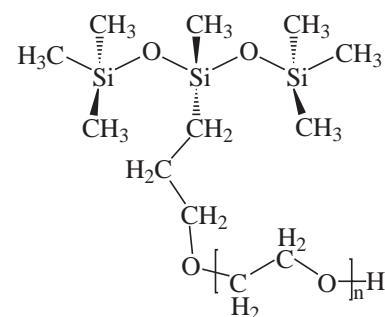
Many other types of copolymers have been made, including graft-types with terminal polar groups, and BAB, AB, (AB)_n, branched and cross-linked structures. Alkyl groups

(such as $C_{16}H_{33}-$), fluorocarbon groups, and combinations of such groups with polar groups may be attached to the siloxane backbone to make materials that are surface active toward a variety of surfaces and interfaces.

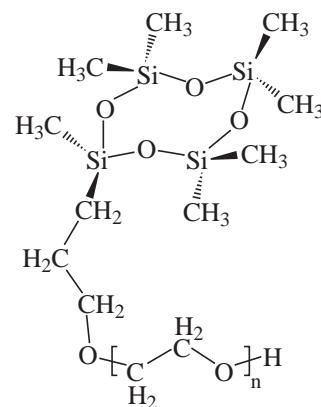
Many small silicone surfactants have also been synthesized and their wetting properties have been characterized. Representative structures are illustrated in the following.



AB, "linear" trisiloxane surfactant, $MDM'E_nOH$.



"Branched" trisiloxane surfactant, $M(D'E_nOH)M$.



Cyclosiloxane surfactant, $D_3D'E_nOH$.

Of all of these possibilities, the wetting properties of the "branched" trisiloxane surfactants, made from 1,1,1,3,5,5,5-heptamethyltrisiloxane, are the basis for the commercial superwetter (or superspread) surfactants. The

wetting properties and the phase behavior of the trisiloxane surfactants will be discussed in the following.

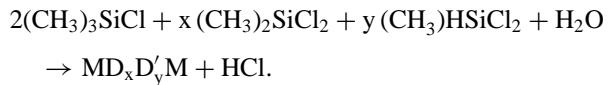
III. SYNTHESIS AND CHEMISTRY

Two routes have been used to prepare silicone polyoxyalkylene surfactants—transesterification and hydrosilylation. In both cases, a reactive silicone and a polyether are first prepared separately, and subsequently coupled together. Certain polar groups are best coupled to silicones by first attaching a small reactive organic group to the silicone and then using the reactive group to attach the polar group.

A. Preparation of Reactive Silicones

The first step in either route is to prepare a siloxane backbone containing reactive sites (either SiOH or SiH) at which to attach the polar groups. There are two ways to do this—cohydrolysis of the appropriate chlorosilanes, or the equilibration reaction.

For example, a siloxane backbone containing SiH groups could be prepared by co-hydrolysis of the appropriate chlorosilanes:



This backbone could also be prepared by (re-)equilibration of the appropriate proportions of end-cap and monomer units as shown in the following reaction:



This reaction is called the equilibration reaction, and is catalyzed by either acid or base. ABA structures, or mixed ABA and rake-type structures can be prepared by substituting M'M' for all or some of the MM in the previously cited reaction. Branched siloxane backbones can be prepared by including T groups. The MDTQ notation is defined in the following table.

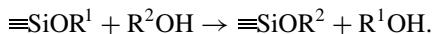
MDTQ Notation for Siloxane Building Block Units

M	$\text{Me}_3\text{SiO}_{1/2}-$	A trimethyl endcap unit
D	$-\text{Me}_2\text{SiO}-$	The basic dimethyl unit
T	$-\text{MeSiO}_{3/2}-$	A three-way branch point unit
Q	$-\text{SiO}_2-$	A four-way branch point unit
M'	$\text{Me}_2(\text{R})\text{SiO}_{1/2}-$	A substituted trifunctional endcap unit
D'	$-\text{Me}(\text{R})\text{SiO}-$	A substituted difunctional unit
T'	$-(\text{R})\text{SiO}_{3/2}-$	A substituted three-way branch point unit

Me	$-\text{CH}_3$
R	H, or (after hydrosilylation) some nonmethyl organic group such as $-\text{CH}_2\text{CH}_2\text{CH}_2(\text{OCH}_2\text{CH}_2)_n\text{OH}$

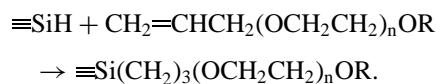
B. Coupling a Polar Group to a Reactive Silicone

The first silicone surfactants were prepared by coupling alkoxydimethylsiloxane polymers and hydroxy-terminated polyoxyalkylenes using the transesterification reaction. An example of this reaction is



R^1 is usually $-\text{CH}_3$ or $-\text{CH}_2\text{CH}_3$, R^2 is a polyalkyleneoxide. The “ \equiv ” denotes that there are three other bonds to the silicon that are not explicitly shown. This reaction yields products in which the hydrophilic groups are linked to the siloxane hydrophobe through an Si—O—C linkage. These materials are useful in polyurethane foam manufacture, and other non-aqueous applications, but in water this linkage hydrolyzes (rapidly at non-neutral pH) to generate silanol and alcohol.

More hydrolytically stable silicone surfactants can be prepared by the hydrosilylation of methyl siloxanes containing Si—H groups with vinyl functional polyoxyalkylenes. An example of this reaction is



This reaction is usually catalyzed using chloroplatinic acid (Speier's catalyst), which can also cause isomerization of the terminal double bond to the less reactive internal position. In order to make sure that all the reactive sites on the siloxane are used, the reaction may be carried out with an excess of allyl polyether.

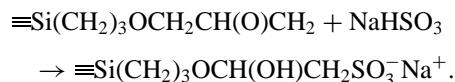
Allyl polyethers with end-capping groups other than —OH are often used. The —OH group also reacts with SiH, albeit slower than a 1,2-vinyl group, so replacing —OH with a nonreactive group prevents this reaction. Examples of such groups include $-\text{OCH}_3$ (methoxy) and $-\text{OC(O)CH}_3$ (acetoxy). A nonreactive end-capping group also prevents a silicone surfactant from reacting with the matrix during polyurethane foam manufacture.

A two-step synthesis is the preferred method of preparing ionic silicone surfactants. First a reactive organic group is attached to the SiH group, then a polar group is attached to that group. This is useful when the reaction conditions of hydrosilylation are incompatible with the desired polar group, although nonionic surfactants have also been prepared by this route. For example, quaternization of halide functional siloxanes leads to cationic silicone surfactants:

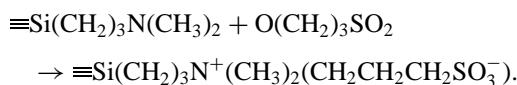
$\equiv\text{Si}(\text{CH}_2)_3\text{X} + \text{NR}_3 \rightarrow \equiv\text{Si}(\text{CH}_2)_3\text{N}^+\text{R}_3\text{X}^-$,
in which X=halide.

C. Coupling Other Groups to Silicones

Anionic silicone surfactants can be prepared by the sulfonation of epoxy-functional silicones as shown in the following:



Zwitterionic silicone surfactants can be prepared by the ring opening of cyclic sultones by aminofunctional silicones as shown in the following:



Increasing the length of the spacer between the siloxane group and the sulfate leads to markedly increased hydrolytic stability while maintaining the surface activity.

The commercial preparation of silicone copolymers containing grafted carbohydrate groups is difficult because the reactivity of the many —OH groups present requires protecting groups. Silicone glycosides for use in skin and hair formulations have been prepared by acid-catalyzed reaction of a saccharide or saccharide source and a dimethicone polyol. Preparation of silicone surfactants with nonionic carbohydrate polar groups generally requires a complex multistep synthesis.

D. Silane Surfactants

Organosilicon surfactants can also be made containing a permethylated Si—C—Si backbone (called a polysilmethylene or carbosilane), or a permethylated Si—Si backbone (called a polysilane). Such structures are not subject to hydrolysis, making them potentially useful in harsher environments. The surface activity of the silane surfactants appears to be inferior to that of the siloxanes.

E. Other Structures

Nonionic, ionic, and zwitterionic derivatives of silicone polyoxyalkylene copolymers have been prepared by reaction with the terminal hydroxyl of the polyether. Use of these derivatives has been claimed for personal care, textile finishing, and building materials applications. Silicone—protein copolymers have been claimed for personal care applications and use in contact lenses. Hydrosilylation can also be used to prepare copolymers in which two or more different types of groups are grafted onto the PDMS backbone. For example, a siloxane backbone

with both alkyl groups and polyoxyalkylene groups is an effective emulsifier for water-in-hydrocarbon oils.

F. Hydrolytic Stability

The hydrolytic stability of silicone surfactants, especially the trisiloxane surfactants, has often been discussed, and is viewed by some as a serious problem. The Si—O—Si linkage is susceptible to hydrolysis in the presence of water:



This equilibrium is catalyzed by acid or base. Rate constants for this reaction are not known but it appears to be quite slow near pH = 7. An induction period is usually observed—only after this time do surface tension and wetting properties markedly deteriorate. Thus, concentrated solutions may appear to be stable for quite long periods of time. Polymeric silicone surfactants are generally more hydrolytically stable. The half-life for dilute M(D'E_{7.5}OMe)M at pH = 7 is more than 40 days. Under strongly acid or basic conditions or at sustained temperatures >70°C hydrolysis leads to rapid loss of surfactancy for most silicone surfactants.

IV. SURFACTANCY

A. Surface and Interfacial Activity

A useful way to compare the surface activity of different surfactants is in terms of their “efficiency” and “effectiveness.” Efficiency is determined by the surfactant concentration required to achieve a certain surface tension, while effectiveness expresses the maximum reduction of the surface tension that can be obtained for that particular surfactant. The patterns of surface activity of silicone surfactants in aqueous solutions are similar to conventional hydrocarbon surfactants. The surface tension decreases with log(concentration) until a densely packed film is formed at the surface. Above this concentration, the surface tension becomes constant. This concentration is called the critical micelle concentration (CMC) or critical aggregation concentration (CAC). For graft-type silicone surfactants, the slope of the curves below the break is smaller than that for the trisiloxane surfactants—due to the larger surface area of the polymer molecules. For polymeric surfactants, the CAC is not a sharp break in the curve, but a gradual transition region—consistent with the broad distribution of molecular species present.

1. Aqueous Surface Activity

Silicone surfactants are very effective surfactants in water—they lower aqueous surface tension values to

20–30 mN/m, while hydrocarbon surfactants usually have surface tensions between 30 and 40 mN/m. Only the perfluoro-surfactants have lower values—between 15 and 20 mN/m. The effectiveness of silicone surfactants is weakly influenced by the nature of the hydrophilic groups. Nonionic and zwitterionic silicone surfactants are the most effective—giving minimum values for the surface tension around 20 mN/m. Ionic silicone surfactants are usually less effective, but still giving surface tension values around 30 mN/m. The surface tensions of solutions of polymeric silicone surfactants increase with the molecular weight of the silicone hydrophobe, presumably because of conformational entropy at the interface. Silicone surfactants containing mixed EO/PO groups tend to be less effective than all-EO examples.

The surface behavior of a homologous series of trisiloxane surfactants $M(D'E_nOH)M$ with $n = 4–20$ shows that the CAC, the surface tension at the CAC, and the area per molecule each vary with molecular structure in a way that was consistent with an “umbrella” model for the shape of the trisiloxane hydrophobe at the air/water interface. The $\log(CAC)$ and the surface tension at the CAC both increased linearly with EO chain length. No studies have been done systematically varying the size of the siloxane hydrophobe.

2. Nonaqueous Surface Activity

The surface activity of silicones in nonaqueous media such as the polyols used in polyurethane foam manufacture results from a combination of the low surface energy of the methyl-rich siloxane species and insolubility determined by molecular weight (see footnote 1). Low-molecular-weight silicones are soluble in many organic solvents. As the MW increases, the solubility decreases. In the MW range in which the polymer is just marginally soluble, it shows notable surface activity, even profoaming. This means that a higher MW silicone surfactant will have a greater tendency to segregate to the surface, even with the same proportions of solvent-loving and solvent-hating groups. In nonaqueous media this can mean the difference between significant surface activity and solubility. Siloxane polyether copolymers lower the surface tension of a variety of organic liquids including mineral oil and several polyols from values of about 25–30 mN/m to values near 21 mN/m.

3. Interfacial Surface Activity

An agent that lowers the air–water surface tension will also lower the interfacial tension between an aqueous solution and most oils. The degree of interfacial tension lowering depends on compatibility between the hydropho-

bic groups of the surfactant and the oil. Interfacial tensions of 2–10 mN/m have been measured for aqueous solutions of silicone surfactants against mineral oil, and smaller values, 0–2 mN/m, against silicone oils. Interfacial tensions between aqueous solutions of a trisiloxane “superwetter” surfactant and normal alkanes vary linearly from 0.025 mN/m for hexane to 0.4 mN/m for hexadecane. These small values indicate that trisiloxane surfactants are very effective at lowering interfacial tensions against low energy hydrocarbon substrates, which is consistent with their ability to cause spreading over such substrates.

B. Aqueous Phase Behavior

Surfactants are useful because they adsorb at interfaces to lower surface and interfacial tensions. Their presence there leads to viscoelastic surface properties and provides a repulsive barrier that stabilizes colloidal particles and foam films. Surfactants also self-associate in solution to form a variety of aggregates ranging from globular, worm-like, and disk-shaped micelles, to bilayer structures such as vesicles. Attractive interactions between the aggregates leads to condensation to form liquid crystal phases. The type of aggregate or liquid crystal phase formed by a particular surfactant, and the progression of liquid crystal phases formed with increasing surfactant concentration, temperature, and salt level can be rationalized using a simple model based on molecular packing. For example, $M(D'E_{12}OH)M$, which has a relatively large polar group, tends to form highly curved aggregates such as globular micelles.

Knowledge of surfactant self-association is important because it controls the rheology and freeze–thaw stability of formulations, and the ability to form and stabilize emulsions and microemulsions. Micelles, vesicles, microemulsions, and liquid crystal phases can all be used as delivery vehicles for perfumes or other active ingredients. Self-assembled microstructures are currently a popular route to templated nanostructured materials.

Two other important aspects of surfactant phase behavior are the cloud point and the physical form of the neat surfactant at ambient temperature. Nonionic silicone surfactants are usually liquids at ambient temperature, while ionic silicone surfactants are waxy solids. Nonionic silicone surfactants containing polyoxyalkylene groups become less soluble in water with increasing temperature, as do nonionic hydrocarbon surfactants. The temperature at which they become insoluble is called the cloud temperature, or cloud point. The general dependence of the cloud point on the weight fraction of polyoxyalkylene and the end-cap of the polyether group is similar to hydrocarbon surfactants. The effects of electrolytes also follow

the usual relationships. Unlike hydrocarbon surfactants, these relationships cannot be simply related to molecular structure using the HLB system although HLB values are sometimes quoted. No simple empirical relationship analogous to the HLB system has been developed yet to predict the emulsifying or cloud point behavior of silicone surfactants.

1. Trisiloxane Surfactants

The aqueous phase behavior of the trisiloxane surfactants in binary mixtures with water has been documented in some detail. These surfactants follow patterns of phase behavior similar to hydrocarbon surfactants—as the polar group decreases in size, there is a progression from globular micelles to worm-like micelles to bilayer structures and then inverse structures. There are two important differences. First, the trisiloxane hydrophobic group is significantly wider and shorter than (for example) a linear C₁₂H₂₅ group, which has about the same degree of hydrophobicity. The length of the trisiloxane group is only 9.7 Å compared with 15 Å for C₁₂H₂₅ while its volume is larger, 530 Å³ compared with 350 Å³ for C₁₂H₂₅. This causes the type of aggregates formed by these surfactants to be shifted toward lower curvature. For instance, C₁₂E₇ forms globular micelles, whereas M(D'E₈OH)M forms bilayer microstructures. Second, the cloud point curve of nonionic trisiloxane surfactants is governed mostly by the size of the polyoxyethylene group. This means that the phase diagram for M(D'E₈OH)M contains a lamellar phase region similar to C₁₂E₄, while the cloud point curve is more similar to C₁₂E₇.

Few studies exist for ionic siloxane surfactants. Several trisiloxane anionic, cationic, and zwitterionic surfactants have been found to form micelles, vesicles, and lamellar liquid crystals. As expected, salt promotes a tendency toward smaller curvature structures.

2. Polymeric Silicone Surfactants

The situation is not so well ordered for polymeric silicone surfactants. Most of the studies that have been published have made use of commercial grade copolymers with varying levels of characterization. There are no commercially available homologous series of polymeric silicone surfactants. Some manufacturers are more willing than others to specify the chemical structures of the products they provide. Even when specific chemical structures are given, undisclosed details of synthesis may give products with the potential to behave quite differently.

The aqueous aggregation behavior that has been reported for graft-type and ABA triblock silicone surfactants differs substantially from that of the block EO/PO

copolymers (the Pluronics®, for example). Those surfactants tend to form globular micelles, and their phase diagrams often contain many small regions of liquid crystal phases, including gel phase and cubic phases. In contrast, polymeric silicone surfactants tend to give simpler phase diagrams with one or two larger regions of liquid crystal phases. Many silicone polyoxyethylene copolymers form bilayer microstructures, including vesicles and lamellar phase liquid crystal. The phase behavior of these copolymers progresses from lamellar to hexagonal phases with increasing proportion of polyoxyethylene. Salts and polyols have the same effect on the liquid crystal phase behavior of silicone surfactants as they do for hydrocarbon polyoxyalkylene-based surfactants. Ternary phase behavior has been reported for a few graft-type silicone polyoxyethylene copolymers. Lamellar and hexagonal phases were found in mixtures with the cyclic silicone oil, D₄, and with hexadecane. A region of water-in-oil microemulsion was found along the surfactant–oil axis for both oils, but the region was quite small for the hydrocarbon oil.

The ability of such complex and polydisperse molecular structures to form ordered structures such as bilayers and liquid crystal phases is attributed to the unusual flexibility of the siloxane backbone. A simple calculation of surfactant shape can be made from the hydrophobe volume, V, the hydrophobe length, L, and the (known) cross-sectional area of the polyether group:

$$S = V/A^* L.$$

A value of S = 1 predicts bilayer aggregates and lamellar phase, while a value of 1/3 predicts spherical micelles and cubic phase. Intermediate values indicate the formation of worm-like micelles and hexagonal phase. Surprisingly, this back-of-the-envelope calculation can be applied to predict the progression of phases formed by polymeric silicone surfactants as well as the trisiloxane surfactants.

Silicone surfactants that incorporate both polyoxyethylene and polyoxypropylene are widely used but very few phase diagrams have been published. Below the cloud temperature, these surfactants form clear isotropic solutions with water at all concentrations. Salt shifts the cloud temperature downward, but does not otherwise change the behavior. No liquid crystal phases have been reported.

Many block copolymers self-assemble in the melt state to form microsegregated ordered phases. Although silicone and polyoxyalkylene are strongly mutually phobic, and therefore would be expected to follow this pattern, the microsegregation behavior of this class of copolymers has not been investigated.

3. Silane Surfactants

The trimethylsilane surfactant, $(\text{CH}_3)_3\text{SiC}_6\text{H}_{12}(\text{OCH}_2\text{CH}_2)_5\text{OCH}_3$ is interesting because it is hydrolytically stable and a good wetting agent. This surfactant forms two separate regions of lamellar phase, L_α . One of the L_α regions, and a region of sponge phase, L_3 , are (unusually) found above the cloud temperature boundary.

C. Ternary Phase Behavior

There is a considerable patent art concerning preparation of transparent mixtures of water and low-molecular-weight silicone oils using a silicone–polyoxyalkylene copolymer as the surfactant. These mixtures are often called microemulsions in the literature in the sense of being transparent mixtures of water, surfactant, and oil—however, they are transparent because of small particle size, or because of index of refraction matching.

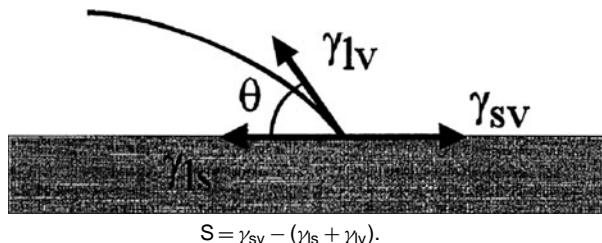
Mixtures of trisiloxane nonionic surfactants with cyclic and short linear silicone oils form extensive regions of microemulsion as well as liquid crystal phases. Trends in phase behavior with EO chain length, salt concentration, and oil molecular weight are similar to the alkyl ethoxylate/alkane/water systems. The phase behavior shifts toward more positive curvature and higher temperatures with increasing EO chain length. Salt shifts microemulsion regions to lower temperature. Higher MW oils require higher concentrations of surfactant to form one-phase microemulsion. The trisiloxane surfactants are efficient surfactants toward silicone oils such as D₄, forming one-phase microemulsion of a 1:1 mixture of water and oil at surfactant concentrations of 6–8 wt%. Trisiloxane surfactants with larger polyoxyethylene groups form hexagonal and cubic liquid crystal phases with cyclic and short linear silicone oils.

Mixtures of several other small silicon-based surfactants, including a silane polyether surfactant, and a silane, a trisiloxane, and a carbosilane amine surfactant with hexamethyldisiloxane (MM) have also been characterized. The polyether surfactants exhibited behavior similar to the C_iE_j surfactants, while the amines required the addition of co-surfactant to form microemulsions. All of the surfactants were relatively inefficient, requiring about 15% surfactant to form one-phase microemulsion of a 1:1 water/oil mixture.

Mixtures of low-molecular-weight silicon-based surfactants and co-surfactants have been used to prepare a self-dispersing microemulsion of silicone agents applied to building materials to impart water repellency. The structure of the surfactants used was not disclosed, but they are described as being themselves reactive so that they bind to the surfaces of the building materials and become part of the water-repellancy treatment.

D. Wetting

Surfactants promote wetting because they have lower surface and interfacial tensions. The energy balance that determines spreading is expressed by the spreading coefficient, S:



Surfactants lower γ_{lv} , and usually γ_{ls} , making S more positive and promoting spreading. The term “wetting” is used in this article to mean spreading to a thin (complete) wetting film with zero contact angle. “Spreading” is a more general term and can describe (for example) relaxation of a drop of liquid on a surface to any final contact angle <90 degrees. Unlike pure liquids, the surface and interfacial tensions of surfactant solutions depend on transport of surfactant to the interface—as the surface is stretched by spreading, surfactant must be transported from the bulk to the expanding interface to maintain low interfacial tension values. In addition, since the area is increasing faster at the spreading edge than near the middle of the drop, a surface tension gradient will develop leading to Marangoni effects. The interlocked roles of fluid mechanics, transport, and Marangoni tractive forces in controlling surfactant enhanced wetting has not yet been quantitatively accounted for.

The ability of silicone surfactants to promote spreading plays an important role in their use in paints and coatings, personal care products, textiles, the oil industry, and as adjuvants for pesticides. The good spreading properties of silicone surfactants have been attributed to low adhesive forces between individual molecules in interfacial films. The time to wet by the Draves wetting test depends on the size of the siloxane hydrophobe and the length of the polyoxyethylene group—the most rapid wetting is observed for surfactants with the shortest siloxane groups and the smaller EO groups.

The unusual ability of the trisiloxane surfactants to promote complete wetting of very hydrophobic surfaces such as polyethylene is called “superwetting” or “superspread-ing.” This wetting ability was first documented in the 1960s when it was noted that aqueous solutions of certain small silicone polyethers rapidly spread to a thin film (complete wetting) on low energy hydrophobic surfaces. It was determined that the best wetting agents are based on small siloxane groups containing 2 to 5 methylsiloxane

units, and that the best wetting was usually observed for surfactants with limited solubility (that is, those that tended to form stable turbid dispersions). The wetting ability has been related to low dynamic surface tensions, the unusual shape of the trisiloxane hydrophobe, the presence of bilayer aggregates (vesicles) in the aqueous phase, and Marangoni effects.

Superwetting has now been shown to share its main characteristics with a variety of ionic and nonionic, micelle- and vesicle-forming hydrocarbon and siloxane surfactants that spread to a wetting film (zero contact angle) on non-water-wettable substrates such as polyethylene. Many common surfactants, such as sodium dodecyl sulfate, DO NOT spread to a wetting film even on relatively hydrophilic substrates (contact angle small but greater than zero). Of those that do, three features appear to be shared in common: (1) there is a maximum in spreading rate as a function of substrate surface energy and (2) surfactant concentration, and (3) there is a critical wetting concentration (CWC) that is significantly higher than the critical aggregation concentration (CAC). Although the surface tension becomes constant at the CAC, the contact angle does not reach zero until a concentration some 3–5X higher.² At least on liquid substrates it has now been demonstrated that the rate of surfactant transport to the interfaces controls the spreading.

Neat liquid trisiloxane surfactants (in common with many amphiphilic materials) spread quite differently on hydrophilic and hydrophobic surfaces—on hydrophobic surfaces the profile at the edge of a droplet of neat liquid trisiloxane surfactant forms steps of approximately the height of a bilayer of the surfactant.

E. Mixtures of Siloxane and Hydrocarbon Surfactants

In spite of its obvious practical importance, the behavior of mixtures of silicone surfactants and hydrocarbon surfactants has not been studied in detail. Early patents claimed generally synergistic behavior for a wide range of combinations of silicone surfactants with hydrocarbon surfactants. Mixed CMC and surface tension results vary from strongly synergistic to antagonistic depending on the ionic character of the hydrocarbon surfactant. Mixtures of the trisiloxane surfactants with anionic surfactants show strongly lowered CMCs and surface tensions, cationics somewhat less so, and nonionics show lower surface tensions, but higher CMCs. Detergency of mixtures shows enhanced oily soil removal, and tolerance to water hardness,

²This is in sharp contrast with the much studied problem of wetting on water-wettable substrates such as clean glass. Surfactant solutions spread on such surfaces well below the CAC.

just as for mixtures where both components are hydrocarbon surfactants. Polymeric nonionic silicone surfactants show similar but less pronounced effects. Recent patents claim enhanced wetting and foaming for certain combinations of silicone surfactants with hydrocarbon surfactants.

V. APPLICATIONS

Silicone surfactants are primarily specialty surfactants that are used in applications that demand their special properties. Most applications are based on some combination of their (a) low surface tension, (b) surface activity in nonaqueous media, (c) wetting or spreading, (d) low friction or tactile properties, (e) ability to deliver silicone in a water-soluble (or dispersible) form, (f) polymeric nature, or (g) low toxicity. The major applications have been reviewed in the literature and will be discussed briefly in following sections.

A. Polyurethane Foam Manufacture

The stabilization of foam during manufacture of polyurethane foam (PUF) was the first, and remains the largest single commercial application of silicone surfactants. PUF is formed by the reaction of polyols and isocyanates. Polyurethane foams range from rigid pneumatic resins to flexible porous elastomers. Rigid PUF is used primarily as an insulating material in construction, piping, and packaging. Flexible PUF is used as a cushioning material in furniture, bedding, carpet underlay, automobiles, and packaging. A number of thorough reviews of polyurethane foam processing are available. The silicone surfactant functions to emulsify the mixture of incompatible materials, stabilize the blowing foam, keep urea particles from aggregating, and govern film rupture for open-cell foams. The process is extremely complex, and difficult to study realistically. Mixtures of silicone surfactants are frequently used to achieve properties that cannot be obtained with single materials. The ability of silicone surfactants to stabilize and control the drainage of the thin films separating growing bubbles appears to be a critical property.

B. Inks, Paints, and Coatings

Surface active materials perform many functions in inks, paints and coatings. The principle uses of silicone surfactants in inks and coatings include defoaming, deaerating, improved substrate wetting, and enhanced slip properties. Although polydimethylsiloxane is used to control foam in many applications, its use in water-based coatings tends to cause formation of defects such as fish-eyes and orange-peel. Early use of silicone polyether copolymers also

experienced difficulties with residual levels of silicone homopolymer that led to defects. However, the purity of currently available silicone surfactants for use in coatings has eliminated this problem. The composition of silicone copolymers for use in coatings includes alkyl, aryl, fluorinated, and polyester functional groups as well as dimethyl siloxane and polyether groups. The particular functionality can be tailored to optimize compatibility with the particular formulation under consideration.

Two types of coating defects are caused by air entrainment—craters (caused by macrofoam) and pin-holes (caused by microfoam). In solvent-based coatings these problems can be dealt with using either silicone fluids or silicone polyether surfactants. In water-based coatings, the materials of choice are the surfactants. The mechanisms of these two problems are different. Macrofoam involves relatively large bubbles, and is best dealt with by promoting rapid rupture (before the coating cures) using a defoaming agent. Relatively insoluble silicone polyethers perform this function best. Microfoam involves very small bubbles that leave pin holes at the coating surface if not eliminated. It is thought that more soluble silicone surfactants cause these small bubbles to rise faster and thereby eliminate the problem. Some silicone surfactants are effective at both functions.

Although it seems obvious that reducing surface and interfacial tensions should be beneficial to spreading of coatings, dynamic interfacial effects and the ability to control surface tension gradients are critical to successful use of surfactants in inks and coatings, particularly for water-based inks and coatings. The usefulness of silicone surfactants as wetting agents is due to their ability to lower surface and interfacial tensions and thus facilitate spreading. Because silicone surfactants can lower surface tensions to values substantially below those of hydrocarbon surfactants, they are effective when hydrocarbon surfactants are not. Although the trisiloxane “superwetting” surfactants are the most effective in promoting wetting, polymeric silicone surfactants are also used for this purpose. As the use of water-based coatings, and plastic engineering materials increases, the need to spread aqueous coatings on low energy substrates will demand increasing use of highly effective wetting agents such as the trisiloxane surfactants.

Silicone surfactants are also used for slip and mar resistance in coatings. These terms imply the ability to impart a durable shiny appearance to a coating that resists abrasion. This property is attributed to segregation of the silicone segments of the copolymer into the surface to provide lubricity. Silicone surfactants with larger silicone groups (which have the most desirable slip and mar resistance properties) may make it difficult to apply a second coat on top of the first.

C. Agricultural Adjuvancy

Increasing use of no-till farming techniques brings with it the need for environmentally friendly herbicides such as glyphosate. Application of such water-soluble herbicides requires the ability to wet waxy weed leaf surfaces to achieve penetration and efficacy. This requirement is the basis of the use of the trisiloxane surfactants as agricultural adjuvants. In the literature of this field silicone surfactants are usually called organosilicone adjuvants. Many different combinations of herbicides, weeds, and adjuvants have been evaluated—certain adjuvants work better on certain weeds with particular herbicides. The interested reader is referred to general references in the bibliography to access this rather large body of literature.

Penetration of the herbicide into the plant leaves is essential to achieve efficacy. The adjuvant promotes wetting of the leaf—and in some manner, also facilitates penetration of the aqueous solution containing the herbicide through the waxy leaf cuticle. Blends of hydrocarbon surfactants with the trisiloxane superwetters are able to achieve much of the required wetting at a lower cost. The hydrolytic stability of the trisiloxane surfactants is a problem in this application because formulations are normally not used at neutral pH. Because of their hydrolytic stability, a group of silane surfactants has recently been introduced to this application.

D. Personal Care

Several types of silicones are used in personal care applications. Oil-soluble silicones and resins are used to promote spreading and film formation by organic oils and waxes. Silicone copolymers with hydrophilic groups are often viewed as a means to overcome the incompatibility of silicone oil with both water and hydrocarbon oils—allowing the technologist to obtain the “dry silky” feel of silicones in a water-dispersible form. Silicone polyethers are used mainly in aqueous formulations such as shampoos and shower gels. In shampoos they improve combing, add gloss, and impart a “dry silky” feel to hair. They have also been shown to reduce the eye irritation of anionic surfactants. Efficacy is determined by solubility—lower HLB versions are more substantive. Incorporation of amine or quaternary nitrogen polar groups improves substantivity in hair-conditioning products. Silicone polyethers can be added to many types of cosmetics and skin care formulations to impart a “nongreasy” feel to the skin. In lotions they impart smoothness and softness to the skin, and defoam, thereby minimizing whitening on rub-out. They have also been found to help prevent cracking in soap and syndet bars.

Microscopic closed capsules called vesicles or liposomes can be prepared from phospholipids and many synthetic surfactants including silicone polyether copolymers. Vesicles are useful as a delivery vehicle for skin care actives. The ability of phospholipids to form closed structures in water consisting of bilayer sheets, called vesicles or liposomes, was discovered many years ago. Since then a multibillion dollar market has emerged for skin care products containing vesicles as a delivery vehicle for skin care actives. The advantages of using a vesicle as a delivery vehicle include enhanced delivery of the active into the upper layers of the skin, controlled release, and wash-fastness. Recently, synthetic surfactant systems have been developed which also form vesicles and avoid some of the problems of natural phospholipid materials. Silicone polyether surfactants with balanced silicone and polyether groups also form vesicles upon dispersion into water that have been shown to encapsulate water-soluble substances, and solubilize lipophilic substances. Silicone vesicles offer improved ease of processing, the ability to combine the aesthetic properties of silicones with effective delivery, and chemical stability.

E. Textiles

Silicone surfactants are used in textile manufacture to facilitate wetting, dispersion of water-insoluble substances, and as spinning and sewing lubricants. They are used as components in lubricants to enable the lubricant to spread quickly and completely even at very low pickup amounts. This property is an example of silicone surfactants lowering the surface tension of organic media. They function as fiber lubricants and release agents. Silicone polyethers are unique in being thermally stable lubricants with good wetting and low coefficients of friction at high speeds. Silicone surfactants are used to prepare silicone fluid-in-water emulsions for deposition onto textiles. They are also used to improve the adhesion of adhesives to silicone-treated fabrics. In certain instances silicone copolymers migrate to the surfaces of fibers during and after extrusion to modify their properties. Obviously, these varied uses involve many different properties of the silicone materials used. Use of silicone surfactants as wetting agents in aqueous textile treatments has been discussed only in terms of laboratory results; little published information is available about actual use.

F. Foam Control

Foam control process aids are the largest single category of process aids used in the chemical industry, and silicone foam control agents are an important segment of this category. The application of silicone fluids compounded

with hydrophobic particulates to control aqueous foams has been extensively discussed and is beyond the scope of this article. Silicone polyethers are used as foam control agents in diesel fuel defoaming, in the manufacture of plastics such as polyvinyl chloride, in polymer dispersions, in inks, paints and coatings, and in some household products. Diesel fuel usually has some moisture in it that strongly influences the action of foam control additives—which must be chosen such that they function over the expected range of moisture contents. The origin of the foaminess is not well understood, but silicone polyethers are effective defoamers as long as they are neither completely soluble in the fuel, nor absorbed and deactivated by the water. Polyoxypropylene containing copolymers appear to be the most effective. Silicone polyethers become less soluble in water with increasing temperature, and copolymers with certain ranges of polyoxyalkylene content have a cloud point. In the vicinity of the cloud point, these copolymers are effective defoamers, although few applications specifically cite this as the controlling mechanism.

G. Emulsification

No systematic studies of the use of silicone surfactants as emulsifiers have yet been published. Silicone polyoxyalkylene copolymers with a relatively high molecular weight and a high proportion of silicone are effective water-in-silicone oil emulsifiers. Similar copolymers with added alkyl groups are effective water-in-hydrocarbon oil emulsifiers. These silicone copolymers form an interfacial film of high viscosity at the oil–water interface that sets up a very effective barrier to droplet coalescence. Characterization of the microstructure of semisolid water-in-oil emulsions formed using terpolymer silicone surfactants shows that the water droplets are dispersed in a gel network of petrolatum with no indication of any ordered structures such as liquid crystals. These copolymers are used to prepare transparent water-in-oil emulsions (often with an active ingredient in the internal phase) for use as deodorants or antiperspirants as well as cosmetics and other personal care products. Their use as drug delivery vehicles has also been claimed. These copolymers can also be used to prepare multiple emulsions not requiring a two-pot process. Such multiple emulsions are claimed to have useful encapsulation and controlled delivery properties for personal care and pharmaceutical applications.

H. Supercritical CO₂

Supercritical CO₂ (scCO₂) is an attractive solvent for cleaning and as a medium for chemical reactions. A number of copolymers have been developed as surfactants

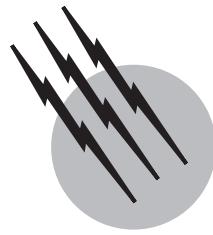
for scCO₂ including some based on polydimethylsiloxane, for example, poly(dimethylsiloxane)-b-poly(methacrylic acid) (PDMS-b-PMA) (MW 5500 g/mol PDMS, 900 g/mol PMA), can be used to stabilize an organic latex in either a nonpolar medium, dense CO₂, or water. Silicone surfactants have been claimed for cleaning applications using scCO₂ as well as stabilizers for dispersion polymerization in scCO₂. The copolymers have been shown to form aggregated structures and lower the interfacial tension between water and scCO₂.

SEE ALSO THE FOLLOWING ARTICLES

COATINGS, COLORANTS, AND PAINTS • LIQUID CRYSTALS
• MICELLES • SURFACE CHEMISTRY • SURFACTANTS, INDUSTRIAL APPLICATIONS

BIBLIOGRAPHY

- Alexandridis, P. (1997). "Poly(ethylene oxide)/poly(propylene oxide) block copolymer surfactants," *Curr. Opin. Colloid Interface Sci.* **2**, 478.
- Artavia, L. D., and Macosko, C. W. (1994). Polyurethane flexible foam formation. In "Low Density Cellular Plastics" (N. C. Hilyard and A. Cunningham, eds.), pp. 22–55, Chapman & Hall, London, UK.
- Bergeron, V., Cooper, P., Fischer, C., Giermanska-Kahn, J., Langevin, D., and Pouchelon, A. (1997). Polydimethylsiloxane (PDMS)-based antifoams. *Colloids and Surfaces A* **122**, 103–120.
- Brook, M. A. (2000). "Silicon in Organic, Organometallic, and Polymer Chemistry," J Wiley, New York.
- Fink, R., and Beckman, E. J. (2000). "Phase behavior of siloxane-based amphiphiles in supercritical carbon dioxide," *J. Supercrit. Fluids* **18**, 101–110.
- Hill, R. M. (ed.) (1999). "Silicone Surfactants," Marcel Dekker, New York, See all chapters in this book.
- Hill, R., and Christiano, S. (1996). Antifoaming Agents. In "Polymeric Materials Encyclopedia" (J. C. Salamone, ed.), p. 285. CRC Press, New York.
- Hill, R. M. (1997). Siloxane Surfactants. In "Specialist Surfactants" (I. D. Robb, ed.) pp. 143–168, Blackie, London.
- Hill, R. M. (1998). "Superspreading," *Curr. Opin. Colloid Interface Sci.* **3**, 247–254.
- Johnston, K. P. (2000). Block copolymers as stabilizers in supercritical fluids. *Curr. Opin. Colloid Interface Sci.* **5**, 351–356.
- Kendrick, T. C., Parbhoo, B., and White, J. W. (1989). Siloxane polymers and copolymers. In "The Chemistry of Organic Silicon Compounds" (S. Patai and Z. Rappaport, eds.), pp. 1289–1361, Wiley, New York.
- Klein, K.-D., Knott, W., and Koerner, G. (1996). Silicone surfactants—development of hydrolytically stable wetting agents. In "Organosilicon Chemistry II, From Molecules to Materials" (N. Auner and J. Weis, eds.), p. 613, VCH, Weinheim.
- Kricheldorf, H. R. (1996). "Silicon in Polymer Synthesis," Springer Verlag, New York.
- Nemeth, Z., Racz, G., and Koczo, K. (1998). "Foam control by silicone polyethers—mechanisms of cloud point antifoaming," *J. Colloid Interface Sci.* **207**, 386–394.
- O'Lenick, A. J., Jr., and Parkinson, J. K. (1996). "Three dimensional HLB," *Cosmetics & Toiletries* **111**, 37.
- Reisch, M. S. (1999). "Paints & Coatings," *Chemical & Engineering News* **77**, 22–33.
- Schlachter, I., and Feldmann-Krane, G. (1998). Silicone surfactants. In "Novel Surfactants" (K. Holmberg, ed.), pp. 201–239, Marcel Dekker, New York.
- Sela, Y., Magdassi, S., and Garti, N. (1995). Release of markers from the inner water phase of W/O/W emulsions stabilized by silicone-based polymeric surfactants. *J. Control. Rel.* **33**, 1–12.
- Wasan, D. T., and Christiano, S. P. (1997). Foams and antifoams: a thin film approach. In "Handbook of Surface and Colloid Chemistry" (K. S. Birdi, ed.), pp. 179–215, CRC Press, New York.



Strategic Materials

John D. Morgan

Industrial College of the Armed Forces

- I. Current U.S. Laws and Agencies
- II. Expanding Supplies
- III. Changing Requirements
- IV. International Concerns

GLOSSARY

Strategic and critical materials Materials required by attackers or defenders in overt or covert hostilities, including declared wars, guerilla wars, insurgencies, sabotage, terrorism, cyberwars, disinformation, and espionage.

WATER and food are daily essentials. Most wars have been fought to gain access to resources. The 41st chapter of Genesis describes how Egypt amassed a huge stockpile of grain 3700 years ago. In recent centuries colonial wars were intended to increase access to needed materials. In the 20th century World War I lasted more than 4 years. Then, counting from Japan's seizure of Manchuria in 1931, World War II lasted 14 years. In both wars expanding land, sea, and air forces required full mobilization of the resources of the major belligerents and use of unprecedented quantities of fuels and materials. [Figure 1](#) illustrates some major applications of materials in military material. Assured supplies of many materials are essential in peace or in war.

I. CURRENT U.S. LAWS AND AGENCIES

The Defense Production Act of 1950, as amended [50 USC, App. 2061 *et seq.* (DPAct)], and The Strategic and Critical Materials Stock Piling Act of 1946, as amended [50 USC, Sect. 98 *et seq.* (SPAct)], establish basic policies and programs. Title I of the DPAct authorizes use of priorities and allocations to direct inadequate supplies to essential uses, and Title III authorizes expansion of supplies. The DPAct designates energy as a strategic and critical material and states that the term materials includes "any raw materials (including minerals, metals, and advanced processed materials), commodities, articles, components (including critical components), products, and items of supply; and any technical information or services ancillary to the use of any such materials, commodities, articles, components, products, or items." The DPAct further states that the term national defense means "programs for military and energy production or construction, military assistance to any foreign nation, stockpiling, space, and any directly related activity" including "emergency preparedness activities conducted pursuant to Title VI of The Robert T. Stafford Disaster Relief and Emergency



FIGURE 1 Strategic and critical materials.

Assistance Act" (42 USCA, No. 5195 *et seq.*). The SPAAct states that the term strategic and critical materials means "materials that would be needed to supply the military, industrial, and essential civilian needs of the United States during a national emergency, and are not found or produced in the United States in sufficient quantities to meet such need." The SPAAct further states that the term national emergency means "a general declaration of emergency with respect to the national defense made by the President or by the Congress." The SPAAct authorizes stockpiling, resource investigations, supply expansions, and research to conserve strategic materials and to develop substitutes and alternates.

By Executive Orders the President has delegated responsibilities under the above acts to agencies with related functions, expertise, and funds. The Departments of Agriculture, Commerce, Energy, and the Interior collect and publish information on resources, production, imports, exports, and uses of materials. These four plus the Department of Defense (DoD), the National Aeronautics and Space Administration, and the National Science Foundation sponsor materials research. The Department of Commerce also supervises export/import controls. The DoD's Defense Logistics Agency (DLA) manages the National Defense Stockpile. The Department of Energy concentrates and stores uranium, deuterium, tritium, and plutonium, develops nuclear weapons, and manages the Strategic Petroleum Reserve (crude oil stored in Gulf of Mexico salt domes). The Department of the Interior's U.S. Geological Survey makes topographic and geologic maps, investigates mineral resources, and collects and publishes information on worldwide minerals, metals, and uses. The Department of the Interior's Bureau of Land Management oversees public lands and manages the government/industry Helium Stockpile in a gastight reservoir near Amarillo, Texas. The Department of Agriculture encourages the production of certain commodities and holds stocks of some. The Treasury Department holds stocks of precious metals for coinage and backing the currency. Working with federal, state, and local governments, the Federal Emergency Management Agency is responsible for emergency planning, preparedness, mitigation, response, and recovery.

II. EXPANDING SUPPLIES

At the end of World War II, Russia, in control of most of the northern part of the Eurasian landmass, became increasingly hostile to the Western democratic nations, initiating the Cold War. In the Far East, China's Red Army extended its control over the mainland. In Southeast Asia returning colonial powers were resisted, impeding a return to prewar

levels of production of raw materials. By 1949 Russia had developed atomic weapons, just 4 years after the United States. Consequently, the United States, Canada, and 10 European nations formed the North Atlantic Treaty Organization (NATO) for mutual security. On June 25, 1950, North Korea attacked U.S.-occupied South Korea. Fears that escalation of the Korean War could lead to World War III increased defense readiness. On September 8, 1950, the DPAct became law. Priorities, allocations, economic controls, and supply expansions soon followed. On October 26, 1950, strengthening earlier arrangements, the United States and Canada agreed that "the production and resources of both countries be used for the best combined results." And in 1950 the United States had begun directly supplying the French in Vietnam, where insurgency had spread. U.S. involvement in Vietnam would increase gradually for 25 years.

By mid-1955,

- the DoD had spent \$110 billion (U.S. billion = 10^9) for hard and soft goods and construction;
- \$8 billion of DPAct financing expanded basic industries, some in foreign nations;
- Rapid Tax Writeoffs under the Internal Revenue Code (26 USC) encouraged private investment of \$37 billion to create new or to enlarge 225 basic industries;
- \$13 billion of Marshall Plan funding expanded Allied essential industries, including strategic materials production;
- production of strategic minerals was further assisted by exemptions from the excess profits tax, and, still in effect, depletion allowances scaled to reflect relative criticality; and
- SPAAct stocks at 270 sites rose to 22 million short tons, valued at \$6 billion.

As a result, communist aggression was halted, and, except for the 1973 Arab Oil Embargo, there have been no major materials shortages. [Table I](#) includes all materials listed under the SPAAct from 1946 to 1999. Most, but not all, were physically stockpiled during the Cold War.

III. CHANGING REQUIREMENTS

Space vehicles, multistage rockets, ballistic, cruise, and guided missiles, supersonic aircraft, remotely piloted vehicles, urban robots, and smart bombs and mines all require strong lightweight metal/ceramic/plastic components and sophisticated electronics with microchips. Observation, aiming, and automatic target recognition rely on infrared and laser devices. Position determination and missile guidance rely on orbiting satellites coupled into

TABLE I All Materials Listed under the Strategic and Critical Materials Stock Piling Act, 1946–1999

Abrasives: Abrasive bauxite, crude aluminum oxide, corundum, emery, silicon carbide, industrial diamonds and diamond dies
Agar
Aluminum: Alumina, aluminum metal, metal-grade bauxite
Antimony
Base metals: Copper, iron ore and iron scrap, lead, tin, zinc
Beryllium: Beryl, beryllium/copper master alloy, <u>beryllium metal</u>
Bismuth
Cadmium
Chromium: Chemical, metallurgical, and refractory ores; <u>chromium metal</u> ; <u>ferrochromium</u> ; ferrosilicochromium
Cobalt
Columbium (niobium): Columbite, columbium metal, columbium carbide powder, ferrocolumbium
Forest products: Balsa, cork, mahogany
Germanium
Indium
Iodine (for chemical uses)
Jewel bearings and natural sapphires and rubies
Leather: Calf and kip skins, cattle hides
Loofa sponges
Magnesium
Manganese: Battery, chemical, and metallurgical ores; <u>ferromanganese</u> ; ferrosilicomanganese; manganese metal
Medicinals: Emetine, hyoscine, iodine, morphine sulfate, opium, quinidine, quinine
Mercury
Molybdenum: Ferromolybdenum, molybdenum disulfide
Monazite and rare earths
Natural fibers: Abaca (manila), extralong-staple cotton, hemp, henequin, hog bristles, jute and burlap, kapok, raw silk and silk waste and noils, sisal, waterfowl feathers and down, wool
Natural oils: Castor, coconut, palm, rapeseed, sesame, sperm, tung
Nickel
Nonmetallic minerals: Asbestos—amosite, chrysotile, and crocidolite; barite, calcite, celestite, cryolite, English chalk, fluorspar—chemical and metallurgical; graphite—amorphous, flake, and lubricating and packing; kyanite, muscovite and phlogopite mica, <u>quartz crystals</u> , <u>refractory bauxite</u> , steatite talc
Optical glass
Pepper
Petroleum and petroleum products
Platinum-group metals: <u>Iridium</u> , osmium, <u>palladium</u> , platinum, rhodium, ruthenium
Pyrethrum
Radium
Rayon fiber (aerospace grade)
Rubber: Crude rubber, latex, guayule seeds and seedlings
Selenium
Shellac
Silver
Tantalum: Tantalite, tantalum carbide powder, tantalum oxide, <u>tantalum metal powder</u> , <u>tantalum ingots</u>
<u>Thorium nitrate</u>
Titanium: Rutile, titanium metal sponge
Tungsten: Tungsten ores, tungsten carbide powder, ferrotungsten, tungsten metal powder
Uranium
Vanadium: Ferrovanadium, vanadium pentoxide
Vegetable tannins: Chestnut, quebracho, wattle
Zirconium: Baddeleyite, zircon

the Global Positioning System (GPS) and the Geographical Information System (GIS). Underwater observation is assisted by multibeam Side Scanning Sonar. Assuring compatibility of communications, ordnance, and material across just the U.S. Army, Navy, Air Force, Marines, and Coast Guard and the many Reserve Components is a major challenge, while the 1999 Kosovo pacification demonstrated that much needs to be done to achieve interoperability among the 19 NATO nations.

U.S. World War II tanks weighed 30 tons; the present MIAI (HA) Abrams Main Battle Tank is up to 70 tons and consumes 6 gal of fuel per mile. Now tanks weighing less than 30 tons are desired to facilitate airlift and mobility over poor terrain. But improved armor penetrators include shaped-charge explosives, projectiles with tungsten carbide or depleted uranium cores, and self-forging tantalum rounds. Lighter bridging equipment is also needed. Mortars are simple effective weapons in the hands of ground troops. The U.S. World War II 81-mm mortar had a separate cast-steel baseplate weighing 48 lb. Now under development are much lighter mortars with composite baseplates and thin steel barrels encased in titanium. A ground force combatant carrying water, food, bandoliers of ammunition, and a dual-caliber automatic rapid-fire weapon with infrared scanning and laser sights, wearing appropriate clothing plus body armor and a helmet with two way TV and multichannel voice radio, and carrying the batteries needed to power the electronics may have a load of up to 100 lb—twice what would be desirable.

The availability of new materials has reduced older requirements for many common materials. U.S. 1950 plastics production was 1 million tons. Now it is about 46 million tons, 40% of the weight of steel but more than three times the volume of steel. Plastics are now widely used in weapons, ammunition, vehicles, ships, boats, construction, containers, and packaging. Synthetic minerals replace natural diamonds, sapphires, rubies, mica, and quartz crystals. Small battery-powered timepieces replace larger, less accurate mechanical ones. Fiber-optic cables and wireless networks speed communications. Synthetics replace cork and balsa in flotation applications, abaca in marine hawsers, hog bristles in brushes, feathers and down in jackets, leather in footwear, and extralong-staple cotton in parachute shrouds. Synthetic rubber replaces natural, although for carrier-based aircraft and huge earthmovers, natural rubber is preferred. Chemicals replace pyrethrins in insecticides and tung oil and shellac in sealants. Research showed that pepper, once considered for stockpiling, was not a preservative in military rations, although pepper is still widely used as a flavoring agent. New uses increase requirements for some materials. Rhenium is now added to chromium/nickel/cobalt jet engine alloys. Titanium is increasingly used in airframes. Once used un-

separated as “mischmetall” (mixed metals), each of the separated 17 rare earth metals has specialized uses. First recovered from natural gas during World War I to serve as a lifting gas in barrage balloons and blimps, helium is now used in shielded-arc welding, deep diving atmospheres, and reactor cooling. Liquefaction of air made possible tonnage oxygen for blast furnaces, nitrogen in portable refrigerators, and argon in welding. Worldwide production of basic commodities and technological advances developing new and improved materials have blunted threats of materials shortages in recent years.

The shift of DoD concern from bulk materials to high-technology materials is illustrated by the DoD's DPAct Title III Program Office's list of year 2000 projects: solid-state power semiconductor switching devices, silicon-on-insulator wafers, silicon carbide substrates, high-purity float zone silicon, flat panel displays, continuously reinforced silicon carbide fiber/titanium metal matrix composites, selectively reinforced silicon carbide whisker/aluminum metal matrix composites, semi-insulating indium phosphide wafers, silicon carbide electronics, and eyewear for protection from laser beams.

Environmental concerns and regulations also influence materials requirements. Formerly widely used poisonous metals and compounds are curtailed or eliminated in traditional uses: lead in gasoline, paints, cable covering, roofing, solders, and containers; cadmium in colors, plating, and plastics; mercury in fungicides and antifouling paints; and cyanide in metallurgy. Curtailment of solid, liquid, and gaseous emissions boosts requirements for corrosion-resistant materials, absorbents, and precipitators. At the same time recycling of metals, paper, and plastics adds to the supplies of basic commodities.

Following the breakup of the USSR and the decline of the Cold War, U.S. stockpile goals for most materials have been sharply reduced or zeroed, reflecting the continued reliance on nuclear deterrence, desire for highly mobile armed forces-in-being with state-of-the-art material, assumed ready access to many foreign sources with only minor shipping losses, and curtailment of production of some civilian goods in an emergency. In the last decade the DLA disposed of stockpiled materials valued in excess of \$3 billion by sales in an orderly manner to avoid market disruption and by transfers of materials to other agencies, e.g., titanium to the Army Tank and Automotive Center and precious metals to the Treasury Department. As of September 30, 1999, in fiscal year 1999 the DLA had disposed of materials valued at \$446 million, leaving in inventory more than 70 commodities valued at \$3375 million, including the 12 underlines in [Table I](#) not authorized for disposal, valued at \$582 million. Despite stockpile disposals, and even if a material had never been listed in [Table I](#), all materials are still covered by the broad DPAct.

IV. INTERNATIONAL CONCERNS

Convertible currencies are highly strategic to all nations, and even well-designed counterfeits are strategic to evildoers. Most materials are of strategic importance to industrialized nations. Some maintain stockpiles or encourage domestic industries to hold larger-than-normal stocks. But most rely on many imports. Even Russia, with a land area greater than that of the United States and Canada combined, and espousing autarky for 70 years, imported many manufactured items and significant quantities of bauxite, alumina, barite, cobalt, fluorspar, molybdenum, tin, and tungsten. Petroleum and gas resources are particularly attractive to investors. Rogue states, insurgents, and guerillas find some materials to be highly strategic—coca and opium for conversion to cocaine, morphine, and heroin and small-volume, high-value drugs easy to smuggle to obtain money or weapons. Even tonnage lots of bulky marijuana are smuggled. Illicit drugs pose serious threats to the stability of developed nations, just as in the 19th century opium imports forced on China led to its collapse. Insurgents and guerillas use primitive artisanal hand mining to recover gold, diamonds, sapphires, and rubies—all high-value, low-volume items. Chemical, biological, and nuclear (CBN) materials are highly strategic to rogue nations, terrorists, and saboteurs. Common chemical agents include ricin, cyanide, mustard gas, phosgene, and explosives. Two common materials—fertilizer-grade ammonium nitrate and fuel oil—when properly mixed, make the powerful explosive ANFO. Common biologic agents include anthrax and smallpox. Any radioactive substances can be used for nefarious purposes. Indeed, just the threat posed by unaccompanied packages possibly containing CBN materials causes urban panics.

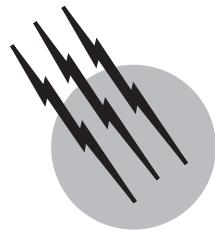
Although industrialized nations have been able to obtain supplies of most materials in recent years and at reasonable prices, there can be no cause for complacency as the world enters the 21st century. Billions of people in less developed nations are now able to use battery-powered small TVs that depict highly exaggerated views of life in the developed nations, and these people are demanding higher standards of living for themselves. As the population surges, pressure on the resources of the earth will increase and the demand for fuels and materials will accelerate. And among the most strategic materials will be water and food.

SEE ALSO THE FOLLOWING ARTICLES

ENERGY RESOURCES AND RESERVES • FUELS • GEOSTATISTICS

BIBLIOGRAPHY

- Department of Defense (annual). "Strategic and Critical Materials Report to the Congress," U.S. Government Printing Office, Washington, DC.
- Energy Information Administration (annual). "Annual Energy Review," U.S. Government Printing Office, Washington, DC.
- Joint Committee on Defense Production, U.S. Congress (1951–1976). "Annual Report of the Activities of the Joint Committee on Defense Production," U.S. Government Printing Office, Washington, DC.
- Morgan, J. D. (1949). "The Domestic Mining Industry of the United States in World War II," U.S. Government Printing Office, Washington, DC.
- National Defense Industrial Association (monthly). *National Defense*, NDIA, Arlington, VA.
- U.S. Geological Survey (annual). "Mineral Commodity Summaries," U.S. Government Printing Office, Washington, DC.



Textile Engineering

Sundaresan Jayaraman

W. W. Carr

L. Howard Olson

Georgia Institute of Technology

- I. Fibers as Components of Engineering Structures
- II. Textile Structures: Production and Applications
- III. Trends in Textile Manufacturing

GLOSSARY

Filling (weft) Yarn that runs widthwise (at right angles to the warp) in a woven fabric; also known as pick.

Jacquard Mechanism invented at the beginning of the nineteenth century by Joseph-Marie Jacquard of Lyons, France; used for weaving elaborate fancy patterns with the help of punched cards carrying binary information. Precursor of the computer punched card.

Nonwoven Textile structure made directly from fibers that is held together by fiber entanglements or by bonding agents (adhesives or fused thermoplastic fibers).

Silver Strand of loose, untwisted fibers held together by interfiber friction. Card sliver when delivered by the carding machine, and drawn sliver when produced by the drawing machine. The fibers in the latter are closer to being parallel (oriented) than those in the former.

Slub Tiny ball of fibers on the yarn that degrades the appearance of the yarn. Sometimes specifically introduced to produce novelty yarns.

Spinning Process of producing yarns from staple fibers by inserting twist in the fine strand of fibers. In the production of man-made fibers, the extrusion of a spinning solution through a spinneret to form filaments.

Warp Yarn that runs lengthwise in a woven fabric; also referred to as end.

TEXTILE ENGINEERING is defined as the art and science of designing and creating structures from *fibers*, the fundamental textile element, and encompasses the design of processes for the production of these structures. The assembled structures have a unique combination of properties: high strength and low bending rigidity. The designing of textile processes and the production of textile structures call for an understanding of two basic scientific disciplines—physics and chemistry—and several engineering disciplines—mechanical, civil, electrical, and chemical. Thus, textile engineering can be regarded as a multidisciplinary applied field of science and engineering. The scope of textile structures has been broadened to include structures that combine high strength with high bending rigidity. These new structures, known as *composites*, find increasing applications in the fabrication of aircraft, helicopters, and space structures. Besides other textile fibers, metallic and ceramic fibers are used as the basic building blocks in the production of these structures.

I. FIBERS AS COMPONENTS OF ENGINEERING STRUCTURES

From an engineering standpoint, textile fibers can be regarded as fine slender rods. They have the unique combination of characteristics of flexibility, because of their fine diameters, and high intrinsic strength per unit weight. Their length-to-thickness ratio, known as the *aspect ratio*, is very high. Fibers of long length are known as *filaments*, whereas short fibers (inches or fractions of an inch long) are called *staple* fibers. For example, a typical cotton fiber is 1-in. long and has an aspect ratio of 1500, whereas a wool fiber, 3 in. in length, has an aspect ratio of 3000. For this reason, fibers are regarded as one-dimensional structures.

A. Classification of Fibers

Traditionally, fibers found in nature (e.g., cotton, wool, silk) have been used to meet one human basic need, namely, clothing. However, with advances in science and technology, newer fibers are being produced, and their share in the production of textile structures is growing. Thus, textile fibers can be classified according to their origin—natural or man-made. These major groups can be subdivided as shown in [Table I](#).

Commercially, cotton is the most important textile fiber. It grows from the seed of the cotton plant, which belongs to the genus *Gossypium*. Silk is extruded by the silkworm in the form of a fine continuous filament, making it the only naturally occurring filament. Wool is obtained from

TABLE I Classification of Textile Fibers

Origin	Type	Source or component	Product
Natural	Vegetable	Seed	Cotton
		Bast	Flax
		Leaf	Sisal
		Fruit	Coir
	Animal	Excreted	Silk
		Hair	Wool
Man-made	Regenerated		Asbestos
		Protein	Casein
		Cellulose	Viscose
	Synthetic	Polyamide	Nylon
		Polyester	Dacron
		Polyacrylonitrile	Orlon
		Polyurethane	Lycra
		Polyolefin	Polypropylene
	Others	Carbon	Thornel
		Glass	Fiberglas
		Ceramic	Fiber FP

the fibrous covering of the sheep. Because of the high cost, the use of silk and wool is fairly limited.

Man-made fibers are classified as either regenerated or synthetic. Regenerated fibers are produced from polymers occurring in nature (e.g., viscose rayon from the cellulose in wood pulp), whereas synthetic fibers are produced by chemical synthesis (e.g., polyester from terephthalic acid and ethylene glycol). The polymer fluid is extruded through a spinneret to form a single continuous filament. Although most man-made fibers are produced as continuous filaments, they are also used as staple fibers. Sometimes several hundred filaments (500–2000) are gathered in a ropelike structure known as a *tow*. A similar structure formed from staple fibers is referred to as *top* (e.g., a wool top).

B. Properties of Textile Fibers

The performance and end-use applications of textile structures are greatly influenced by the properties of the fundamental unit of the structure, the fiber. [Table II](#) provides a summary of the important characteristics of some typical textile fibers.

The fineness or linear density (mass/unit length) of a fiber is expressed in terms of *tex*, which has units of grams per kilometer. Typical fineness values are in the 0.1- to 1-tex range. Fiber strength and extensibility are essential for textile applications. Tenacity, expressed in grams per tex, is a measure of fiber strength, whereas breaking extension characterizes the extensibility of fibers. The extensibility of man-made fibers can be altered during the manufacturing process. For polyester it typically varies between 15 and 55%, whereas for spandex, used in swimwear, it is between 400 and 500%.

Moisture regain is defined as the ratio of the mass of absorbed water in a specimen to the mass of the dry specimen. Natural and regenerated man-made fibers are generally hydrophilic and have fairly high regains (e.g., cotton has a regain of 7%). On the other hand, most synthetic fibers are hydrophobic and have low regains (e.g., polyester has a regain of 0.4%). The moisture regain of fibers influences the wearing comfort of garments and the processing of fibers into yarns. So, cotton garments are more comfortable than those made of polyester. Polyester is highly prone to static, which causes problems in spinning. To minimize static problems, fiber producers apply antistatic finishes to man-made fibers.

Fibers are affected by light, microorganisms, and chemicals, and the degree of degradation depends on the fiber type. Cotton, for example, is attacked by mildew and fungi. Fibers are generally not affected by mild alkalis and acids. However, strong acids attack cotton, and strong alkalis can damage polyester.

TABLE II Properties of Textile Fibers

Fiber	Diameter range (μm)	Density (g/cm^3)	Tenacity ^a (g/tex)	Breaking extension ^a (%)	Moisture regain ^a (%)	Attacked by strong
Cotton	11–20	1.52	40	7	7	Acids
Wool	20–40	1.31	11.5	42	14	Alkalis
Silk	10–15	1.34	40	23.4	10	Alkalis
Viscose rayon	≥ 12	1.46–1.54	20	20	13	Acids
Nylon 6.6	≥ 14	1.14	35–60	15–60	4.5	Oxidizing agents
Polyester	≥ 12	1.34	25–50	15–50	0.4	Alkalis
Orlon	≥ 12	1.16	20–30	20–30	1.5	Alkalis
Glass	≥ 5	2.54	76	3	0	Unaffected

^a At 65% relative humidity and 70°F.

II. TEXTILE STRUCTURES: PRODUCTION AND APPLICATIONS

Conventional textile structures combine strength with a high degree of flexibility, and this combination sets them apart from other engineering structures. Flexibility in other engineering structures can be achieved only at the cost of a loss of strength. Discontinuities in textile structures are both necessary and inevitable. For example, a yarn spun from staple fibers has voids or air space in it (~40% voids), and the entrapped air contributes to the warmth and comfort of textile materials. In contrast, discontinuities or voids in other engineering structures act as stress raisers, which can lead to catastrophic failures. Textile structures are soft and have a fine texture, whereas other engineering structures are hard and smooth.

A. Order and Textile Structures

The fundamental goal of the different operations in the production of textile structures is the imposition of order on the mass of fibers. In the simplest of structures, *yarn*, the slender one-dimensional fibers are assembled in such a way that not only is the fiber flexibility preserved, but the resultant interlocking of the fibers provides the necessary strength to resist large-scale deformations. In essence, yarns are a linearly ordered one-dimensional assembly of fibers. Yarns are interlaced or interlooped to form two-dimensional planar structures known as *fabrics*. A typical fabric will have $\sim 10^{10}$ fibers per square yard, and so the task of the textile engineer is to ensure the correct selection and ordering of the fibers and yarns to produce a useful structure. A variation of this traditional process is the direct conversion of fiber to fabric, in which the fibers are “laid” and bonded to produce two-dimensional structures known as nonwovens. The terms *one-* and *two-*

dimensional refer to the macroscopic perception of the structure and its dimensions.

B. One-Dimensional Textile Structures

A classification of one-dimensional textile structures is shown in Table III. Yarns can be directly extruded as continuous filaments, or they can be spun from staple fibers. The different yarn types are also listed in the table. A yarn is regarded as a one-dimensional structure because its transverse dimensions are very small compared with its length.

1. Yarn

In the traditional yarn manufacturing process, a stock of fibers is opened, cleaned, and attenuated to form a fairly parallel strand of fibers. This linear assembly of fibers has no coherence, and the fibers can easily slide past one another. Coherence, and hence strength, are imparted to the fiber strand either by increasing the interfiber frictional forces or by “bonding” them together using an adhesive.

TABLE III One-Dimensional Textile Structures

Textile structure	Type	Product
Yarns	Continuous	Monofilament
		Multifilament
	Spun (staple fiber)	Homogeneous
		Blended
Others		Textured filament
		Plied (multifold)
		Combination
		Novelty
Braids		
Ropes		
Cords		

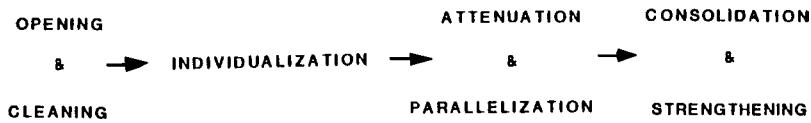


FIGURE 1 Fiber to yarn: sequence of operations.

The sequence of operations on the fibers is shown in Fig. 1. The opening and cleaning machines loosen up the fibers and, in the case of cotton, remove trash, dust, and other impurities in the fibers. Fiber individualization is brought about by the carding machine, in which the fibers are subjected to a brushing action between two metallic wire surfaces moving at vastly different speeds. The fine web of fibers delivered by the carding machine is condensed into a ropelike form known as a *sliver*. The linear density of a typical sliver is in the 3- to 4-ktex range, and there are thousands of fibers in the cross section.

The yarn, in contrast, has ~ 100 fibers in its cross section, and its linear density is in the 5- to 50-tex range. To reduce the linear density of the sliver, the next step in the sequence is the attenuation process, also known as *drafting* or *drawing*. Drafting is carried out by passing the sliver through two pairs of rollers rotating at different speeds (Fig. 2). The first pair, through which the sliver is fed, rotates slower than the second pair, and this results in the reduction of fibers in the cross section of the material delivered by the second pair. The ratio of surface speeds is called the *draft*, and the reduction in linear density of the sliver is proportional to the draft. Drawing, while involving draft, is in fact an intermediate process used for evening out linear density variations and parallelizing fibers.

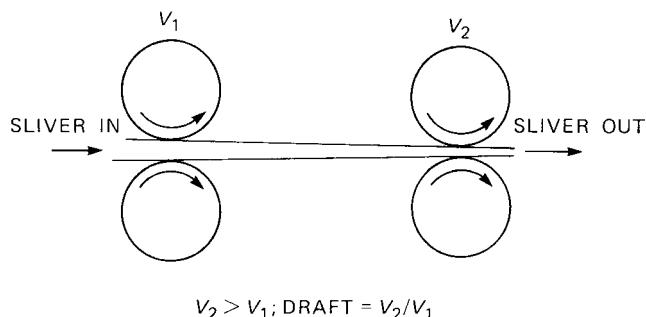
There are three principal techniques for imparting coherence and strength to the strand of fibers attenuated to the desired linear density. They are *twisting*, *wrapping*, and *bonding*.

a. Twisting. Twisting is the most commonly used technique for fiber consolidation, and *ring spinning* incor-

porates this principle. Here, the drafted strand emerging from the nip of the front rollers is led down through the *traveler*, which is rotating at high speeds (~ 40 m/s) on the inside flange of a ring and inserting twist (see Fig. 3). The takeup package, or *bobbin*, drags the traveler, and the difference in speed between the two causes the strand, the yarn, to wind on the package. The typical production rate is 20 m/min per spinning position. Twist can be inserted in one of two directions: clockwise (properly known as S twist) or anticlockwise (Z twist), as shown in Fig. 4. The production rates in ring spinning are limited by the fact that the yarn package has to be rotated at very high speeds ($\sim 16,000$ rpm), causing traveler burnouts and excessive aerodynamic and centrifugal stresses on the yarn.

Open-end spinning utilizes the principle of twisting for yarn formation but eliminates the need to rotate the package. Here, the attenuated strand of fibers is individualized, deposited on a rotating surface, and then reassembled into a yarn at a free growing end. Hence, the name *open-end spinning* is used. In rotor spinning, the widely used open-end method, the fibers are fed into the circumferential groove in a drum rotating at high speeds (up to 80,000 rpm). The rotating free end of the yarn peels the fibers from the rotor surface, and the yarn is withdrawn simultaneously and wound onto the takeup package. The production rate is 100–120 m/min per spinning position.

b. Wrapping. Coherence and strength can be imparted to the fiber strand if it is wrapped with a continuous filament yarn (Fig. 5). The resulting wrapped yarn has no twist in the core and is bulky. In a variation of the wrapping method, the staple fibers themselves wrap around the core of fibers. The drafted strand of fibers passes through

FIGURE 2 Pair of drafting rollers (V_i = velocity). Note the reduction in strand thickness.

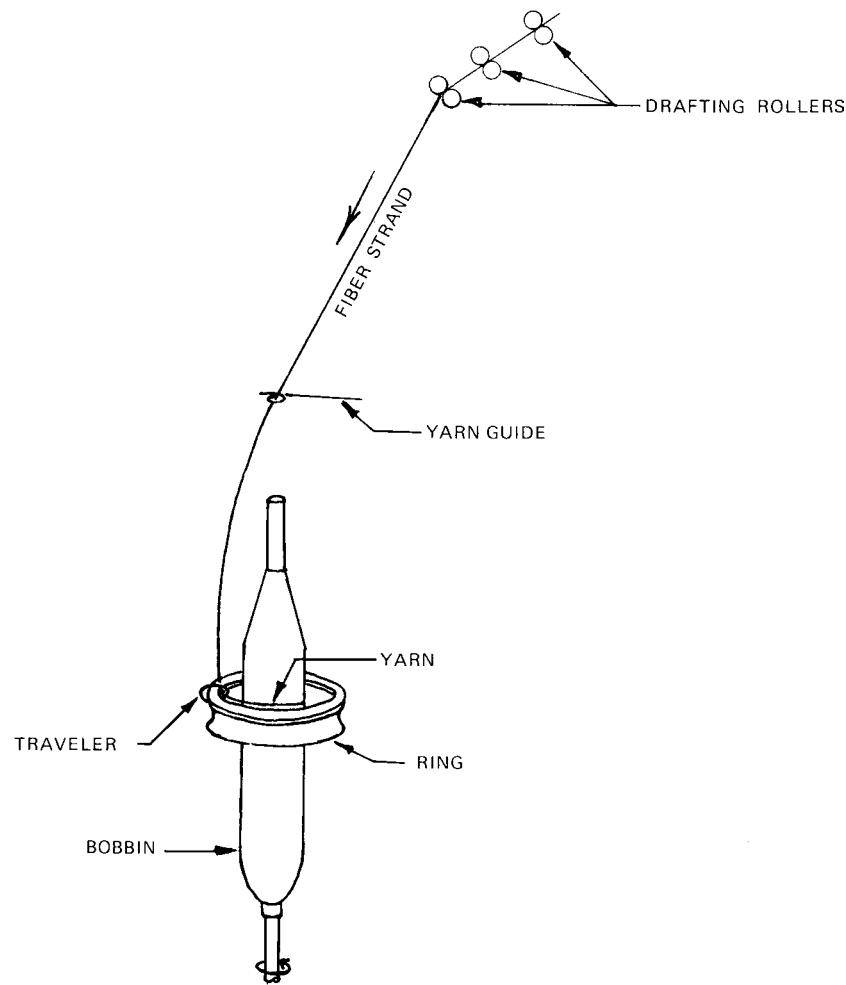


FIGURE 3 Material flow in ring spinning.

two successive twisting units that insert twist in opposite directions. As a result, the twist in the strand emerging from the second unit is “zero”; however, some of the fibers (those at the edges) that are not twisted in the first unit wrap around the core in the second and consolidate it. Thus, the yarn consists of a fairly parallel core of fibers held together by the wrapper fibers. This principle is also known as the “false-twist” method because there is no real twist in the yarn. In the commercial form, air-jet spinning, jets of air are used at the twisting zones. The production rate is ~ 180 m/min per spinning position.

c. Bonding. The fibers in the drafted strand can be bound together by an adhesive to form a yarn, and this is the principle used in the Twilo twistless spinning method. The yarn, although behaving like a solid rod, can withstand the rigors of fabric production, after which the adhesive is washed off to yield a fabric with a high degree of bulk.

2. Blended and Composite Yarns

Blended yarns are produced by combining two or more staple fibers, usually in the drawing process. The purpose of blending is to produce a yarn that advantageously

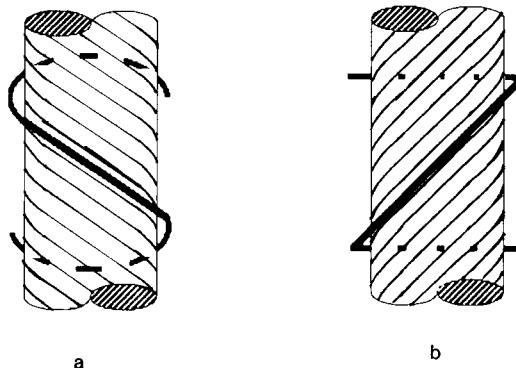


FIGURE 4 Twist in yarns. (a) S twist; (b) Z twist.

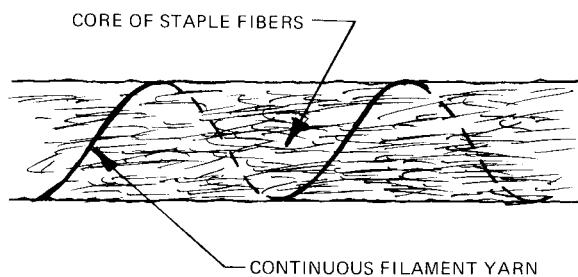


FIGURE 5 Yarn formation by wrapping.

combines the properties of the component fibers. For example, in a common blend like polyester–cotton, cotton provides the comfort properties, whereas polyester gives strength to the blend.

A variation of blended yarns is *core-spun yarn*, which has a continuous filament core and a staple fiber sheath. In a composite yarn, called the Bobtex ICS yarn, a filament yarn is fed through a molten polymer onto which a layer of staple fibers is subsequently applied. Very high production speeds of 1000 m/min are possible.

3. Plied and Novelty Yarns

Plied yarns are produced by twisting together two or more singles yarns. Generally, during plying, the yarns are twisted in the direction opposite to the singles-yarn twist to give a torque-balanced yarn. The balancing of twist such that the finished structure will exhibit little tendency to coil or kink on itself is done by monitoring the *twist multiplier* used in the singles yarn and in the plies. Twist multiplier is defined as the ratio of turns per inch to the square root of the cotton count. For maximum strength, the twist multiplier used in the ply structure should equal the twist multiplier of the singles yarn. For balance, the twist multiplier of the plied structure should be $\sim 80\%$ of the singles-yarn twist multiplier.

Yarns with special effects on their surfaces are known as *novelty yarns*. They are produced by varying the linear density across the yarn and also by introducing *slubs*.

4. Continuous Filament Yarns

The production of monofilament yarns is the same as man-made fiber production. Several monofilament yarns are combined to yield a multifilament yarn. Twist is inserted to hold the filaments together and to prevent abrasion and breakage of individual filaments. Continuous filament yarns have a smoother surface than staple fiber yarns. Thus, fabrics made from these yarns tend to be shiny and less bulky and also lack the warmth and wearing comfort associated with fabrics made of staple fiber yarns. The fiber crimp and the discontinuities (voids) in the staple fiber yarns account for the bulk and comfort of these fabrics.

Texturing is a process that modifies the structure and geometric characteristics of continuous filament yarns to impart properties normally associated with staple fiber yarns. In texturing, the continuous filament yarn is deformed (by twisting), and this deformation is set (by heating and then cooling). As a result, a bulky structure is obtained because the set deformations prevent the filaments from lying in a parallel order.

5. Yarn Structure and Properties

The linear density of staple fiber yarns is often expressed by an “indirect” system of measurement in which the length of yarn per unit weight is specified. Denoted as the *cotton count*, it is expressed as the number of hanks per pound, where a *hank* is equal to 840 yd. Thus, a 30s Yarn means that 25,200 ($= 30 \times 840$) yd of the yarn weigh 1 lb. An individual yarn is referred to as a *singles* yarn; note that the *s* is used here and with indirect yarn numbers.

The geometry of the twisted yarn is shown in Fig. 6. During the yarn formation process, the fibers on the surface follow longer paths (in comparison with the fibers in the core), develop higher tensions, and displace the core fibers that are under low tension. This cyclic interchange, known as migration of fibers, produces a self-locking, yet flexible structure with a fairly regular geometry. The amount of twist inserted is determined by the fiber length and diameter, the linear density of the yarn, and the intended end use. A highly twisted yarn is hard and compact and tends to snarl, whereas a low-twist yarn is bulky and soft.

The important yarn properties are strength, elongation, appearance, and uniformity. Strength and elongation are influenced by the amount of twist inserted; appearance depends on the fineness of the fibers used and the carding process. Neps (small groups of fibers) can be introduced due to improper carding. Uniformity or variation in linear density is greatly influenced by the drafting process. A yarn with excessive linear density variation is unsuitable for most applications. An exception is novelty yarns, which are produced by varying the linear density along the length of the yarn.

6. Ropes

A rope is a collection of twisted yarns built in stages. Initially, several singles or multi-ply yarns are twisted into strands, and several of these strands are twisted in the reverse direction to yield a rope. To achieve a compact structure, proper tension should be maintained in the final twisting operation. Nylon, because of its high tensile strength, good abrasion, and flexural properties, is the most commonly used fiber for making ropes and cords. Polyester is used in applications where a stiffer or less elastic rope is needed. Large ropes are formed by plying

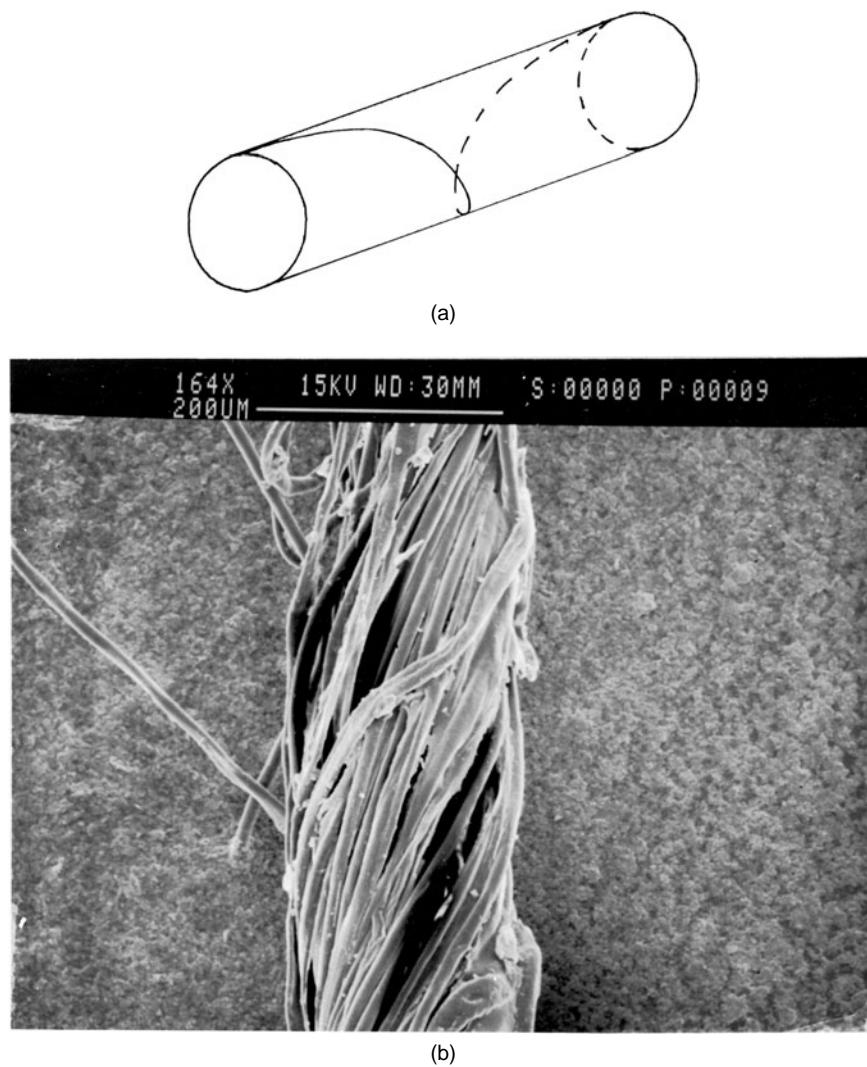


FIGURE 6 Yarn structure. (a) Twist geometry; (b) scanning electron micrograph of yarn.

additional layers of strands or smaller plied structures on top of a core.

The surface character of a rope is determined by the twist organization of the individual strands and plies of these strands in constructing the rope. A *cord* structure uses alternating twist organization. For example, a ZSZ cord is formed from Z-twisted strands (yarns) that are plied with S twist to form an intermediate structure. These are then plied with Z twist to form the final structure. Cord structures are softer than *hauser* (or *hawser*) structures, which use twist on twist in their construction. A hauser may, for example, have a ZZS twist organization.

7. Braids

A braid is essentially another one-dimensional textile structure. However, when braiding is done over a form or shaped object, as in the formation of a structural el-

ement for composite material structures (e.g., a rocket nose cone), it ceases to be a one-dimensional structure. Rather than being simply wrapped in layers as may occur with rope formation, a braid involves interlacing simultaneously two opposing sets of strands to form a tube or flat tape. One or more threads or yarns are wound as a strand onto a braiding tube. A set of 16, 22, or more tubes is placed in carriers, which are driven around a circular track with each half of the set rotating in opposite directions. As carriers meet, they alternately deviate from the center of the circular path, inward or outward, to avoid collision.

As strands are pulled off during this process, they are interlaced. The common interlacing patterns are (1) *diamond* braid (1/1), which interlaces at each strand; (2) *regular* braid (2/2), which interlaces at every other strand and is the most common braided structure; and (3) *hercules* braid (3/3), which interlaces at every third strand.

Braids are characterized by very easy extension until the strands touch or jam against one another. A rope made by braiding will have less tendency to rotate when subjected to a tensile force than a rope made by twisting. Braiding is a slower process than twisting. Braids are used alone (e.g., shoe laces and sutures) as a protective or binder covering for a core strand and in decorative applications.

C. Two-Dimensional Textile Structures

Two-dimensional textile structures, generally referred to as fabrics, are characterized by a substantial area and relatively small thickness. Weaving and knitting are the two conventional ways of producing fabrics, whereas the production of nonwoven fabrics represents a nonconventional process.

1. Weaving and Woven Structures

The oldest method of fabric production is *weaving*, in which one set of yarns (*warp*) that runs along the length of the fabric is interlaced with another set (*filling* or *weft*) that runs across the width of the fabric. The ends of the fabric are called *selvedges*. The production of woven structures is carried out on a *loom*. While the sheet of warp yarns runs lengthwise through the loom, the filling yarn is laid singly—one length at a time.

Weaving Preparatory Processes. Analogous to the intermediate steps in the conversion of fiber to yarn, the processes in the conversion of yarn to woven fabric are winding, warping, slashing, and weaving. The purpose of winding is to convert small ring bobbins to larger packages called *cones* or *cheeses* and to remove defects in the yarn such as slubs and thick and thin places. In warping, ends from several hundred of these packages are wound side by side (under uniform tension) on a beam. The length of yarn and the width of the sheet are determined by the length and width of cloth to be woven.

During the weaving process, the warp yarns are under tension and are also subjected to the abrasive action of some of the loom parts (drop wires, *heddles*, *reed*). At the same time, there is interyarn rubbing, and the surface fibers from one yarn become entangled with the fibers on the neighboring yarns. Both situations lead to yarn breakages. The purpose of *slashing* or *sizing* is to improve the weavability of the yarns to minimize warp breaks on the loom. Yarn strength is enhanced by the use of adhesives, and their covering action entraps the protruding surface fibers, resulting in a smooth yarn. Lubricants are used to improve the pliability of the yarns. Several warp beams (giving the required number of ends in the fabric) are combined into a single sheet and led into a size bath containing adhesives, lubricants, and a solvent. Commonly used adhesives are polyvinyl alcohol and all types of starches. Wa-

ter is the general solvent, whereas mineral and vegetable oils, animal fats, and mineral waxes serve as lubricants. The sized yarns are dried on cylinders and wound onto a beam known as the weaver's beam.

a. The weaving cycle. The three primary motions on a loom are *shedding*, *picking*, and *beating up*. During shedding, some of the warp ends are raised and others are lowered according to the weave pattern to create a shed for the filling. The filling is inserted during picking, and the *pick* is then beaten up to form the cloth. In a conventional or *shuttle* loom, the filling yarn is released from a shuttle that traverses from one side of the loom to the other. Weaving speeds are measured in terms of the number of picks inserted per minute and weft insertion rate, which is equal to the product of the picks per minute and the loom width. The typical speed for a shuttle loom 1.1 m wide is 210 picks/min, giving a weft insertion rate of 231 m/min.

The large mass of the shuttle (~1 lb) and the need to accelerate and decelerate it over short distances have been factors that limit production rates. Other methods of weft insertion, in which the shuttle has been replaced, are gaining in popularity. In a projectile or gripper loom, a small gripper carries one pick length of yarn from one end to the other. Fluid jets (air and water) are also used for propelling the weft across the shed. Water-jet looms are generally used for continuous filament yarns; they cannot be used for yarns from hydrophilic fibers (like cotton) because they absorb water. In a rapier loom, metallic arms at either end of the loom are used for picking. The filling is brought from one end of the loom to the center by one arm, where it is transferred to the second rapier, which carries it to the other end, completing the weft insertion process. Typical weaving production rates are listed in Table IV.

b. Woven fabric structures. There are a wide variety of ways in which the warp and filling yarns can be interlaced. The resulting structure influences the functional and aesthetic properties of the fabric. There are three weaves, *plain*, *twill*, and *satin*, from which all other woven structures are derived. The simplest of weave is the plain weave. As shown in Fig. 7a, the first warp thread is over the first pick, under the second, over the third, and so on, whereas

TABLE IV Weaving Production Rates

Loom type	Width (m)	Speed (picks/min)	Weft insertion rate (m/min)
Shuttle	0.9–1.1	180–220	230
Projectile	2.2–5.4	380–420	1100
Air jet	1.25–3.3	900–1200	1900
Water jet	1.25–2.1	1000–1500	2100
Rapier	1.65–3.6	475–500	1200

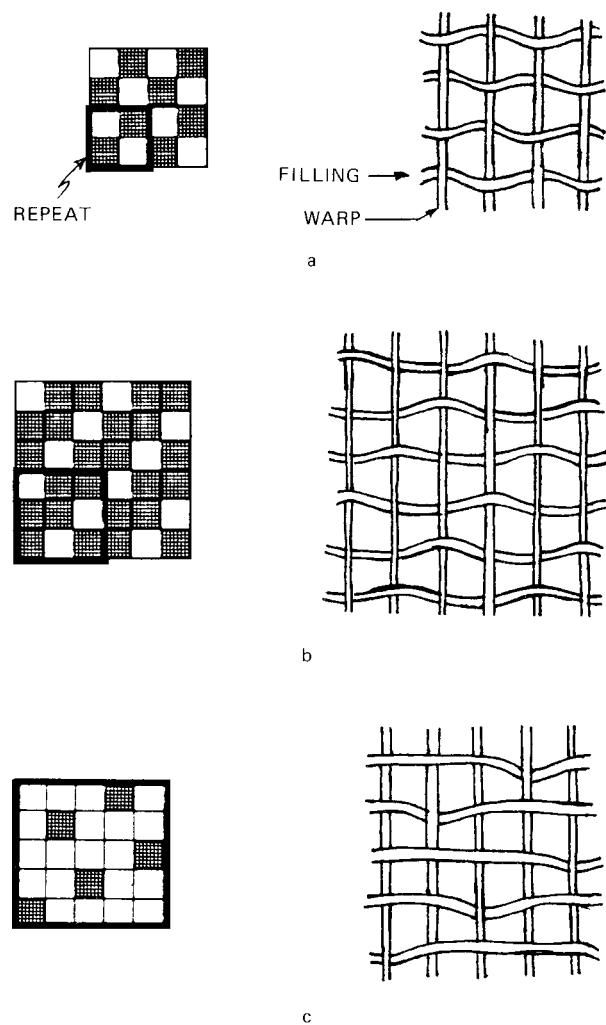


FIGURE 7 Three primary weaves and corresponding interlace-ments. (a) Plain weave; (b) twill weave; (c) satin weave (weft faced).

the second warp thread is under the first pick, over the second, under the third, and so on. The plain weave repeats on two ends and two picks. The plain weave can be extended along the warp or filling or both directions to yield different structures.

The twill weave is characterized by diagonal lines, known as twill lines, that run at angles to the fabric. A simple 2/1 (read 2 up, 1 down) twill is shown in Fig. 7b. The steps due to the staggering of the interlacements produce the twill line. A five-end weft satin (known as sateen) is shown in Fig. 7c. The unique characteristic of the satin weave is that there are no adjacent interlacements in a repeat. In addition, they have long “floats” in either the warp or filling direction. When a low-twist bulky yarn is used for the floating thread, a soft, lustrous fabric with a luxurious feel is produced. For producing simple patterns that repeat over 36 ends and picks, a *dobby* shedding mo-

tion is used. For more elaborate patterns that repeat over several hundred ends and picks, individual control of ends is obtained using the *Jacquard*.

Other woven structures include pile fabrics (terry towels) produced using two sets of warp ends, triaxial fabrics woven from three sets of threads, and double cloths in which two layers of fabric are “stitched” at specific intervals.

2. Knitting and Knitted Structures

Knits are two-dimensional structures produced from yarns by forming loops in the yarn and passing successive lines of loops through one another. Whereas two sets of threads (warp and weft) are used in weaving, knitted fabrics are produced from either a warp set of yarns or a weft yarn. Hence, knitting is divided into two major areas: *weft knitting* and *warp knitting*. The term *interlooping*, in contrast to *interlacing*, describes the knitting process. Figure 8 shows the structures of weft and warp knit fabrics.

a. Weft knitting. Weft knit fabrics are produced predominantly on circular knitting machines. The simplest of the two major weft knitting machines is a jersey machine. Generally, the terms *circular knit* and *plain knit* refer to jersey goods. The loops are formed by knitting needles

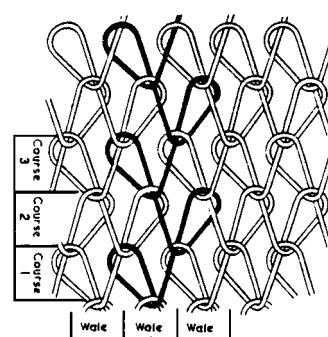
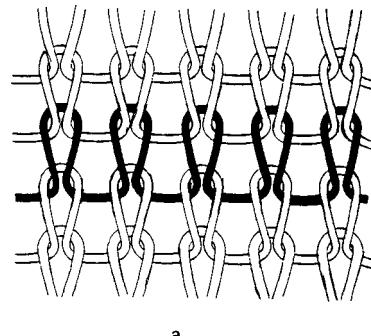


FIGURE 8 Knitted structures. (a) Weft knit; (b) warp knit.

and the jersey machine has one set of needles. Typical fabrics are hosiery, T-shirts, and sweaters.

Rib knitting machines have a second set of needles at approximately right angles to the set found in a jersey machine. They are used for the production of double-knit fabrics. In weft knits, design effects can be produced by altering needle movements to form *tuck* and *miss* stitches for texture and color patterns, respectively. Instead of a single yarn, several yarns can be used in the production of these structures. This increases the design possibilities.

b. Warp knitting. Warp knitting involves a flat needle bed of from 6 to 15 ft in width. There are two major types of warp knitting machines: *tricot* and *raschel*. The structures of the two machines are much more functionally similar than is found in comparing jersey with rib. The warp knitting machine is somewhat more demanding mechanically to operate in terms of precision and input yarn quality demands. The tricot machine is designed to produce fabrics of generally simple design, but at a very high production rate. In contrast, the raschel machine is very versatile, producing fabrics ranging from fine laces to heavy upholstery.

Warp knits use yarn guides for each independent warp yarn beam set to deliver the yarns to the needles in a predetermined pattern. Passing yarn over the needle in warp knitting is referred to as *lapping*. The overlap across the hook of the needle can form an open or closed loop. The underlap passes to adjacent needles. Both the overlap and underlap affect the dimensionality stability of warp knit fabrics.

The two commonly used types of knitting needles are shown in Fig. 9. Latch needles are used on weft and raschel machines, whereas spring beard needles are used on tricot machines.

3. Nonwoven Structures

In a nonconventional method of producing two-dimensional structures, the intermediate (and expensive) step of yarn formation and their geometric ordering are

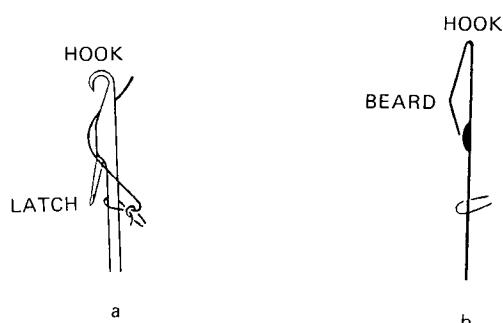


FIGURE 9 Knitting needles. (a) Latch needle; (b) spring beard needle.

eliminated, and fibers are directly converted to a fabric. Unlike a woven or knitted fabric, in which a well-controlled and definable arrangement of yarns prevails, a nonwoven structure usually consists of a web of fibers (in a stochastic arrangement) bonded together by one of several methods.

The first step in the production of nonwovens is the formation of the web. There are three types of webs: parallel-laid, cross-laid, and random-laid. When staple fibers are used, parallel-laid webs are formed by laying a number of carded webs on top of one another in the same direction, giving a highly anisotropic web. Consequently, the properties of the resulting fabric are significantly different in the two principal directions (known as *machine* and *cross* directions). To reduce this difference, a cross-laid web is produced by cross-lapping the webs. For a more isotropic web, special machines are used to produce a random arrangement of fibers, giving random-laid webs. When continuous filament yarns are used for the production of nonwovens, they are directly extruded onto a moving belt. Historically, web formation began as a wet laying process used by papermakers. The dry lay process (described earlier) predominates now.

The formed web can be bonded in one of two ways, either *adhesively* or *mechanically*. In adhesive bonding, a *thermoplastic* fiber or powder that has been blended with the web melts due to the application of heat and binds the fibers together. Alternatively, when the web is produced from continuous filament yarns, they are self-bonded and such fabrics are known as spun-bonded fabrics. Instead of a thermoplastic fiber or powder, adhesive binders such as acrylic resins are used. The web is saturated with the binder and then dried. Print bonding is the selective application of binders to the web in a desired pattern. The fibers have a greater flexibility of movement (since the proportion of binders is small), and the fabric is softer than the one produced by saturation bonding.

Mechanical bonding has its origins in the production of *felt* from wool fibers. Felts are formed by mechanical agitation of wool fibers in a wet state, which causes fiber entanglement and yields a coherent and strong structure. Finishing felt hats by hand rubbing with mercury led to the expression "as mad as a hatter." In *needle punching*, a form of mechanical bonding, barbed needles penetrate a web, causing the fibers from the surface to move *into* the structure. The embedded fibers are trapped by the neighboring fibers, giving a mechanically interlocking structure. Needle punching is usually carried out on both sides of the fabric. Since the fabrics generally tend to be weak, support yarns are used. In another variation of mechanical bonding, known as *stitch bonding*, yarns are knitted (usually warp knit) through the fibrous web. Arachne and Malimo are two commercially available stitch-bonded nonwoven fabrics.

D. Fabric Finishing

The fabric as it comes out of the loom or knitting machine is known as a *gray* or *greige* fabric. It is rough and stiff (due to the size) and does not absorb water, a necessary precondition for dyeing. So the fabric is subjected to a finishing process that enhances its aesthetic and functional properties. Termed *wet processing*, the typical sequence includes *desizing* for removal of the size added during slashing (not done for knitted fabrics since the yarns are not sized), *scouring* to remove surface impurities and improve the fiber's water absorption characteristics, and *bleaching*. *Heat setting* is carried out for fabrics from man-made fibers.

The fabric is then dyed to the desired shade using the class of dye appropriate to the fiber. For example, *reactive*, *vat*, and *direct* dyes are commonly used for cotton, whereas *disperse* dyes are used for polyester. *Printing*, or localized dyeing, is the selective application of dyestuff to specific areas of the fabric to produce a desired pattern. Specific chemical finishes are applied to the fabric to impart some properties such as water and soil resistance, flame retardance, and crease recovery. Dyeing, normally a wet process, can also be carried out as a dry process.

E. Fabric Properties

The aesthetic and functional characteristics of fabrics are influenced by the finishing process, the fabric structure, and the type of fiber. The major functional properties of fabrics are strength, elongation, and resistance to tear and wrinkling. Among the surface properties of interest are resistance to abrasion, *pilling* (tendency to form fiber beads due to abrasion), and fabric handle. Air and water permeability are the major transfer properties of textile structures.

The capacity of a textile fabric to "fall" when it is hung is called *drape*, and this characteristic sets it apart from other two-dimensional structures (e.g., a sheet of paper or a metal plate). Drape is a direct consequence of the fabric's capacity to undergo shear deformation and is related to the stiffness of the fabric. Chemical finishes are used to modify the drapability of fabrics. The assessment of fabric handle (softness, drape, feel) is a purely subjective process, and attempts are being made to link the mechanical properties of fabrics (bending, tensile, shear, etc.) to the handle of fabrics. The Kawabata hand measuring system is one such commercial method.

III. TRENDS IN TEXTILE MANUFACTURING

The Jacquard card used to carry binary information into the weaving process was the precursor of the punched card for the computer and, in fact, was the inspiration for

the birth of the computer card. The field of textile engineering, which triggered the computer revolution, is quite appropriately taking advantage of computers in its various activities.

A. Computers in Manufacturing

One of the primary objectives of a textile manufacturer is to produce a defect-free end product, or at least minimize the occurrence of defects. However, due to the stochastic nature of some textile processes (e.g., yarn production), precise control is not possible and consequently defects are introduced in the material. Drafting or drawing is one such operation in yarn manufacturing whereby linear density variations are introduced in the sliver. To minimize these variations, *autolevelers* operating on either the *feedback* or the *feedforward* principle are made an integral part of drawing machines. The speed of either the input or the output rollers (see Fig. 2) is changed, causing a change in the draft ration and eventually resulting in a change (correction) in the linear density of the sliver.

1. Integrated Production and Quality Monitoring

Modern spinning machines, like the Murata air-jet spinning system, incorporate on-line measurement of yarn evenness; several spinning machines are linked to a central computer, enabling the operator to monitor the operation from the console. Downtime due to stoppages is computed, and machine efficiency is displayed. When the yarn package is built to the desired size, the doffing mechanism removes the package and transports it on a belt to one end of the machine, where it is picked up by another arm and placed inside a container. When the desired number of packages has been placed in the container, it is automatically boxed and rendered ready for shipping. Thus, the role of the human operator, and hence the chance of error, are minimized in the production process.

2. Sizing and Weaving

Microprocessors are also used on the slashing machine to produce a high-quality sized yarn. Parameters such as the viscosity, temperature, and pickup are monitored and regulated by these controllers. In the weaving plant, looms are monitored continuously by individual microprocessors, and the data on the type and duration of stoppages are fed into a central computer. These data are analyzed to spot trends in loom functioning and aid in the early detection of problem looms. The use of this information in dynamically assigning operators to stopped looms is presently being researched and may find its way to the weaving floor in the future. Material handling is being automated with the use of self-guided vehicles.

3. Design and Dyeing

Design workstations with high-resolution graphics terminals and color palettes enable the designer to create and view different structures before actually producing the fabric. The computer can transfer the pattern instructions to the knitting machine for production. In the area of dyeing and printing, the Millitron process from Milliken & Co. represents the state of the art in ink-jet printing. Here, programmed jets spray dyestuff in different colors and patterns onto the material. Since the patterns and colors can be varied easily, small lots can be produced with a minimum of setup time and costs.

4. Fabric Inspection

Inspection of the finished fabric for defects (e.g., broken or missing ends and picks) has long been a slow and tedious process requiring the services of an experienced operator. Defects are classified as major or minor depending on the severity. If the allowable number of defects is exceeded, the fabric is classified as *seconds* and sold at a lower price. Electronic inspection systems are replacing the human eye, making high inspection speeds possible. The fabric "ticket" (a paper containing the fabric particulars that accompanies the fabric) now incorporates not only

the number of defects, but also the place of occurrence of each defect in the fabric so that the garment manufacturer can take this into account while cutting the fabric.

B. Computers in Engineering Design of Textile Structures

Apart from the traditional area of accounting and costing, computers, and especially personal computers with the huge volume of useful software, are being increasingly used in the areas of engineering design of textile structures, production calculations, inventory management, and production planning and scheduling. Figure 10 illustrates the system of constraints in the design of a woven fabric. To arrive at the necessary fabric specifications, the textile engineer must have the flexibility to analyze the effect of the different parameters (counts of warp and filling, number of warp ends and filling picks, the crimp in the warp and filling) on fabric areal density (weight per unit area). Software packages with the ability to manipulate equations are used for this purpose.

C. Textile Operations in the Future

The proliferation of inexpensive computer hardware and software will lead to the wide-spread use of

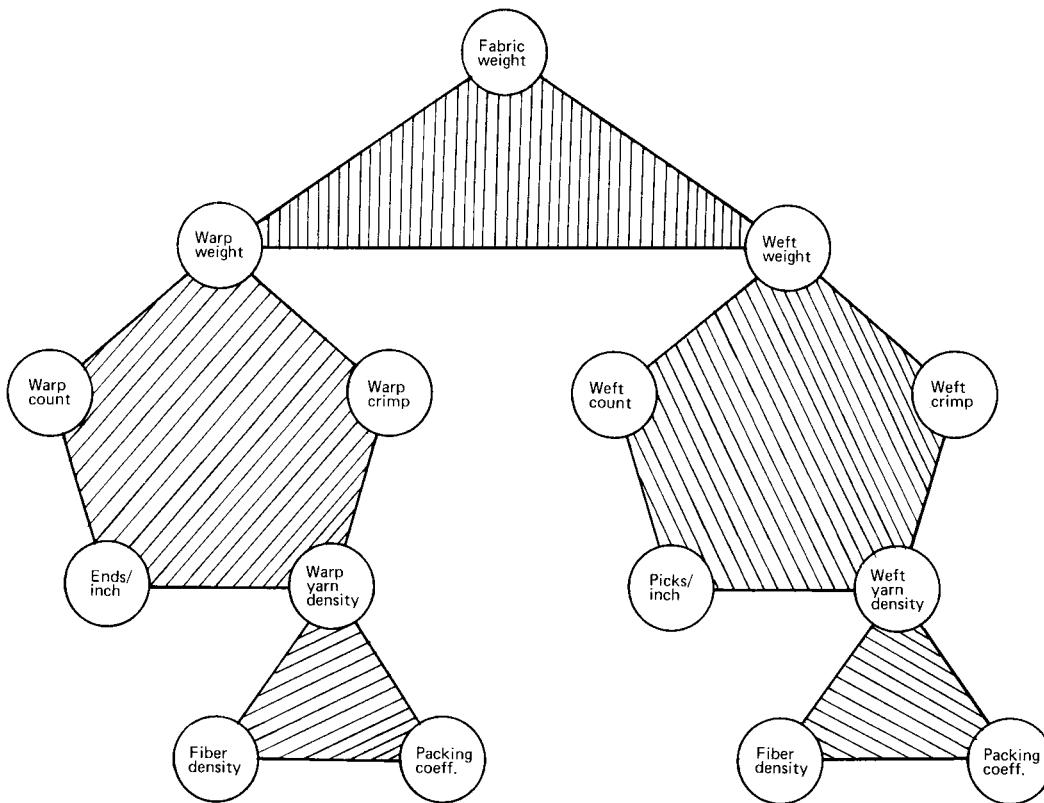


FIGURE 10 Engineering design of woven fabrics: system of constraints.

microprocessors for better control of processes. Research on knowledge-based expert systems in specific areas of textile engineering is in progress. The use of computer networks for communication within a plant and between plants situated in various places will increase. The centralization of operations such as purchasing and research and development is gaining ground. Computers and automation should become an integral part of textile operations for the continued growth of the industry.

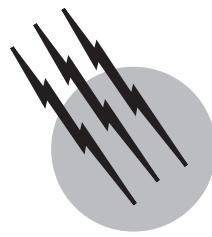
SEE ALSO THE FOLLOWING ARTICLES

BIOPOLYMERS • COMPOSITE MATERIALS

BIBLIOGRAPHY

Acar, M. (1995). "Mechatronic Design in Textile Engineering," Kluwer Academic, Dordrecht/Norwell, Massachusetts.

- Ash, M., and Ash, I. (2000). "Handbook of Textile Processing Chemicals," Synapse Information Resources, Inc., Endicott, New York.
- Goswami, B. C., Martindale, J. G., and Scardino, F. L. (1977). "Textile Yarns: Technology, Structure, and Applications," Wiley (Interscience), New York.
- Joseph, M. (1981). "Introductory Textile Science," 4th ed. Holt, New York.
- Lord, P. R., and Mohamed, M. H. (1982). "Weaving: Conversion of Yarn to Fabric," 2nd ed. Merrow, Durham, United Kingdom.
- Luneschloss, J., and Albrecht, W. (1985). "Nonwoven Bonded Fabrics," Halsted Press, Chichester, United Kingdom.
- Marks, R., and Robinson, A. T. C. (1976). "Principles of Weaving," The Textile Institute, Manchester, United Kingdom.
- Raheel, M. (1996). "Modern Textile Characterization Method," Dekker, New York.
- Spencer, D. J. (1989). "Knitting Technology," 2nd Ed. Technomic, Lancaster, Pennsylvania.
- Vigo, T. L. (1994). "Textile Processing and Properties: Preparation, Dyeing, Finishing and Performance," Elsevier, Amsterdam/New York.
- Vincenti, R. (1993). "Elsevier's Textile Dictionary: In English, German, French, Italian and Spanish—CD-ROM," Elsevier, Amsterdam/New York.



Tin and Tin Alloys

William B. Hampshire

Tin Research Institute, Inc.

I. Mining and Refining of Tin

II. Marketing of Tin

III. Properties of Tin

IV. Applications of Tin

GLOSSARY

Alluvial Describing a tin ore deposit where the primary (lode) deposit of tin ore has been eroded by water.

Base box Measure of tin plate surface area corresponding to 112 sheets, each 20 in. by 14 in. Tin plate gauges and tin coating thicknesses are still referred to in pounds per base box.

Bronze Alloy of copper with tin developed in prehistoric times; used for its corrosion and wear resistance.

Eutectic Two or more metallic elements at a specific composition that melts at a constant temperature, as does a pure metal. For example, an alloy of ~62% tin and 38% lead melts at 183°C, below the melting point of either tin or lead.

Fire refining Production of refined metal by a series of furnace treatments, as opposed to electrolytic refining, which is a selective dissolution-electroplating sequence.

Fusible alloy Metal alloy, usually containing at least some tin, that melts at a low temperature (183°C or below) and is used primarily because of this low melting temperature.

Pewter In the modern sense, a tin-based alloy with anti-

mony and copper additions. It is prized for its decorative appeal.

Placer Another term for an alluvial deposit.

Terneplate Mild steel base with a coating of lead-tin alloy. The tin content of the coating typically is 8–20%.

Tin plate Mild steel base with a coating of pure tin, the material from which tin cans are made.

White metals Family of tin-based (or lead-based) bearing metals, so named for their characteristic color. In the broader sense of the term, pewter is a white metal as well.

TIN is a soft, ductile, metallic element used in prehistoric times to create bronze by alloying the tin into copper. Elemental tin possesses atomic number 50, atomic weight 118.69, and chemical symbol Sn (derived from the Latin name *stannum*). Tin is in the periodic table subgroup with carbon, silicon, germanium, and lead and has some properties similar to the properties of these other elements. The applications of tin typically take advantage of the metal's low melting point, excellent corrosion resistance, and nontoxicity to living organisms. The largest of these applications are tin plate, solder, and tin chemicals.

I. MINING AND REFINING OF TIN

The practice of tin-extractive metallurgy began before recorded history. Since tin has never been known to occur as native metal, it is generally believed that Bronze Age humans obtained bronze by smelting together native copper with a tin ore, undoubtedly the tin oxide mineral called cassiterite.

Even today cassiterite is the dominant ore constituent for commercial exploitation. In some areas of the world, it still exists as primary lode deposits, but most of the production involves secondary or placer deposits, where the original lode body has been eroded by natural forces.

A. Alluvial Mining

In tropical areas of the world the secondary tin deposits are often mined by alluvial techniques. Where possible, dredges, which are large, floating processing plants, are constructed on natural or manmade lakes or, more recently, at offshore locations. Here the dredges use a bucket line or suction cutter to lift the ore from up to 50 m below the water level into the dredge, where a series of jigs, shaking tables, and flotation cells effect beneficiation of the ore, up to a level as high as 70% tin from perhaps 0.02% initially. The overburden from which the tin mineral has been removed is discharged from the rear of the dredge to settle back to the lake bottom.

Obviously the mining of such low-grade ores requires the handling of large amounts of material, and the larger dredges often are capable of processing more than 1000 tons/hr. The separation of cassiterite is energy intensive aboard a dredge, so a high throughput is necessary to justify the high capital cost of the dredge and the high operating costs.

B. Above-Water Placer Mining

Other placer deposits are found where flooding and dredging may not be economical or even possible, and these are typically worked by gravel pump mining. Typically, a pit is created and enlarged as high-pressure streams of water are directed against the sides. This water washes the tin-bearing material to a sump, where the gravel pump lifts the slurry to a processing area.

Normally, the ore is processed by pumping into a long sluice, where the cassiterite settles to the bottom and the lighter non-tin-bearing material is carried away. The partially concentrated ore can then be removed and taken to a concentrator, where jigs, shaking tables, and flotation cells are used to produce a product suitable for smelting.

It has been estimated that ~40% of the world's tin is produced in this manner. Often the operation is a modest one, perhaps run by a single family, where the miner sells

the ore to a custom concentrator, there being relatively little vertical integration in the tin mining industry.

C. Hard-Rock Mining

Primary lode deposits are found in Bolivia and Australia, for example, and these are worked by underground mining techniques. Shafts are sunk and tunnels driven from them into the lodes so that the ores can be removed. These ores are often complex in composition, requiring more expensive processes to extract the metal values.

The mining in Cornwall, England, is actually hard-rock mining, but in fact some tunnels run out under the sea. These Cornish tin mines date back centuries to Roman times. In recent years there has been renewed interest in mining there, although hard-rock mining is seldom cost competitive with other methods.

D. Refining Techniques

Often concentrates are treated before smelting to remove some of the metallic impurities. The exact nature of the treatment depends on the impurities present, but often roasting or leaching processes figure prominently. Roasting drives off certain impurities that are volatile and may convert other impurities, especially if roasted under a controlled atmosphere, to more leachable compounds. Water or acid leaching can then remove additional impurities, but of course the concentrate must then be dried before smelting.

Tin smelting can be carried out in blast furnaces, reverberatory furnaces, or electric furnaces. Regardless of the furnace used, the principles of tin smelting remain the same. The concentrate is mixed with carbon, usually as powdered anthracite, which serves as the reducing agent to change the tin oxide to metallic tin. Also added is limestone, which acts as a fluxing material.

This mixture is charged into the furnace, where it is heated to ~1400°C for ~12 hr with agitation of the bath to promote separation of the impurities from the metal. The furnace is tapped into settling pots, where the metal again sinks to the bottom and slag overflows the top. The metal can then be cast into solid form for additional refining, but even the slag must be reprocessed.

Unfortunately for tin smelters, tin oxides easily combine with silica, which is readily available in tin ores. The resulting tin silicates enter the slag in such quantity that the first slag is too valuable to be discarded. Often a separate furnace is employed, with considerable smelting experience, to remove enough tin from the first slag to leave a second slag for discarding and more impure tin metal for further refining.

Fire refining is preferred whenever the impurity content of the metal allows it. Typically, liquating is a first step and

utilizes a sloped-floor furnace, heated to just above the melting point of tin. The tin-rich metallic phase melts out and runs down the slope, leaving iron, copper, and other higher melting-point impurities behind. Often liquating is followed by a poling process similar to that used in the copper industry. Poles of green wood are immersed into the metal and bring impurities to the surface, where they can be skimmed off. The resulting metal is typically ~99.8% tin.

Electrolytic refining of tin is practiced when certain impurities are present and cannot be removed by fire refining. Again as is the case for copper, impure metal is cast into anodes, which are hung in cells of electrolyte. An electrical current causes the anodes to dissolve and the tin plates out on cathodes hung nearby. The cathodes are made from electrolytically refined tin initially and, when fully plated, are removed and melted and cast into ingot form for shipping. They typically reach a purity of 99.99% or better.

II. MARKETING OF TIN

Throughout the history of its use, tin has been a strategic material. There is evidence to suggest that long-range contacts between peoples during the Bronze Age were actually fostered by a search for tin reserves. Even today, tin is a geographically significant metal, usually produced as a primary metal a long distance from its consumption. **Table I** lists the major tin-producing countries. The major tin consumers are shown in **Table II**.

TABLE I Western World Tin Production by Country (tonnes)

Country ^a	1995		1996	
	Refined tin	Tin in concentrates	Refined tin	Tin in concentrates
Malaysia	41,000	6402	38,100	5175
Indonesia	44,200	46,058	48,900	51,024
Thailand	8,200	1784	11,000	1299
Brazil	17,000	19,500	17,600	19,200
Bolivia	18,000	16,300	16,700	14,802
Peru	—	22,020	500	27,002
Australia	—	7750	—	8496
Portugal	—	4616	—	4625
U.K.	—	2000	—	2000
Africa	—	3200	—	3200
Rest of West	14,000	2000	13,000	2000
Total	142,400	131,630	145,800	138,823

^a China and Russia are major tin producers, but production data are not currently verifiable.

Note: Data are from *Tin Monitor*, published by CRU International, Ltd., London.

TABLE II Western World Consumption of Refined Tin by Country (tonnes)

Country	1995	1996
USA	35,000	36,300
Japan	28,100	26,900
Germany	19,000	19,600
U.K.	10,400	10,600
France	8200	8100
Rest of Europe	23,700	21,300
Other industrialized countries	8000	6000
Developing countries	56,500	58,000
Total	188,900	186,800

Note: Data are from *Tin Monitor*, published by CRU International, Ltd., London.

A. International Tin Agreements

The International Tin Agreement was established primarily to prevent the price fluctuations that plagued the market for many years. These price variations hurt consumers and producers alike, and both groups sought a better balance of supply and demand. The operating body for the tin agreements, currently for the Sixth International Tin Agreement, is the International Tin Research Institute (ITRI) Ltd., established in 1956. Representatives of the member countries meet to discuss the tin industry and tin markets and, when necessary, to vote to take action that will bring about a shift in the supply–demand relationship.

The nature of the ITRI is often misunderstood. It is not a “producer cartel,” because both producers and consumers are represented. The producer countries are apportioned 1000 votes based on their production, and the consumer countries are apportioned 1000 votes based on their consumption, so neither group of countries by itself can force action on the other.

The primary tool available to the ITRI in influencing supply and demand is a buffer stock of tin metal or cash. The buffer stock manager buys or sells tin depending on whether the price is too rapidly declining or escalating, respectively. The exact circumstances under which the buffer stock manager must act, may act, or must not act are defined by the ITRI on the basis of the information it has on tin production costs, mining capacities, future trends, and other factors.

At the time of this writing, the ITRI has invoked its other, more serious provision for influencing the supply side, namely, export controls. By requiring the producing countries to limit their exports of tin, it is hoped that balance will be restored. This effort is extremely difficult for the producing countries, which depend greatly on tin exports for their foreign trade value.

B. Tin Markets

World prices for tin are often based on the quotations on the London Metal Exchange (LME), where tin has been traded for decades. Contracts on the LME are made for immediate delivery or for delivery in 3 months. The contract prices are used to determine the “closing prices,” and there are two of these: “cash” (for immediate delivery) and “forward” (for delivery in 3 months).

Recently the Kuala Lumpur Tin Market has begun a trade in tin. This market is, of course, based on a different currency and therefore introduces the factor of currency exchange rates into the trading of tin metal. There was already a Penang price established by physical trading in concentrates at the Penang smelters. This price has been used as a guide for trading elsewhere, as have the “offerings,” the number of tons of concentrates traded, an indication of supply on a daily basis.

Tin is also traded on most of the major commodity markets—for example, the COMEX in New York. Tin prices are regularly reviewed in several publications, such as the *American Metal Market* and *Iron Age*.

III. PROPERTIES OF TIN

A. Physical Properties

Table III lists some of the properties of tin. The wide range of applications of tin make use of various properties, especially the low melting point, the capacity of tin to form alloys with many other metals, its excellent corrosion resistance, and its nontoxicity. The importance of each of the properties is described in Section IV.

B. Mechanical Properties

At room temperature tin is already at more than one-half of its absolute melting point (505 K). Therefore, the metal is in the high-temperature regime of mechanical behavior. Creep, recrystallization, and grain growth may all occur readily, so the mechanical strength of tin (**Table IV**) is too low to permit its use as a structural material. For this reason tin is nearly always alloyed with another metal or coated onto a stronger metal to provide support. It is quite fortunate that tin coats other metals so well.

IV. APPLICATIONS OF TIN

A. Tin Plate and Canning

1. Tin Plate Manufacture

Historically the largest (~35% worldwide) and most important use of tin has been for coating mild steel sheet

TABLE III Some Physical Properties of Tin^a

Property	Value
Atomic number	50
Atomic weight	118.69
Valencies	2, 4
Density (kg/m ³)	
β -Tin at 15°C	7.29
α -Tin at 13°C	5.77
Liquid tin at mp	6.97
Melting point (°C)	232
Vapor pressure (mm)	
At 1000 K	7.4×10^{-6}
At 1500 K	0.17
At 2000 K	30.6
At 2500 K	638
Boiling point (°C)	~2270
Latent heat of fusion (kJ/g atom)	7.08
Thermal conductivity at 20°C (W/mK)	65
Linear coefficient of thermal expansion at 0°C (°C ⁻¹)	19.9×10^{-6}
Surface tension at mp (mN/m)	544
Expansion on melting (%)	2.3
Electrical resistivity, β -tin at 20°C, ($\mu\Omega$ cm)	12.6
Critical temperature (°C) and pressure (atm)	3730 and 650

^a Corrosion behavior summary: Tin stays bright in dry air at room temperatures. It is generally oxidized rather slowly in harsher atmospheric conditions. Tin is very resistant to high-purity natural fresh waters and to milk products. Tin is resistant to most dilute acid solutions (not resistant to nitric or sulfuric acids) and is resistant to some stronger acid solutions if they are oxygen free. Tin is resistant to most solvents, oils, and other chemicals, the notable exceptions being chlorine, potassium hydroxide, and sodium hydroxide. Tin is less corrosion resistant at elevated temperatures.

to make the product called tin plate. This material is then fabricated into tin cans to preserve food or for the containment of a wide variety of other products. This process is due to an ingenious materials concept that was developed

TABLE IV Some Mechanical Properties of Tin^a

Property	Value
Tensile strength (N/mm ²)	
At 20°C, 0.4 mm/mm min	14.5
At 100°C, 0.4 mm/mm min	11.0
Young's modulus at 20°C (kN/mm ²)	49.9
Poisson's ratio	0.357
Creep strength, approx. life at 2.3 N/mm ² and 15°C (days)	170
Brinell hardness (HB)	
At 20°C, 10 kg/5 mm × 180 sec	3.9
At 100°C, 10 kg/5 mm × 180 sec	2.3

^a Note: The tensile properties of pure tin and many tin-based alloys are very dependent on the rate at which a load is applied.

in Bavaria in about the fourteenth century. The strength, durability, light weight, and fabricability of iron (later steel) were combined with the corrosion resistance and compatibility of a tin coating to yield an important engineering material for pots and pans, lanterns, boxes, and so on. Then when Appert invented the method of preserving foods by sterilization inside sealed containers, it seemed only natural to use tin plate as the container material.

Modern tin plate is rather different from the original material and, in fact, has evolved greatly since the 1930s. Today, cold-rolled steel in huge coils is fed into a processing line that electrolytically coats the surface with tin. At the entry end coils are welded together end to end to give a continuous feed of steel strip. The strip is passed through a series of tanks that clean, pickle, and rinse the surface.

The tanks that follow are electrolytic cells, where tin is plated on the surface. The strip may be moving at up to 600 m/min. Any one of three electrolytes may be employed for the plating process. Pure tin anodes dissolve under the influence of the applied current and plate out on the strip as it passes. Typically, both sides of the strip are plated, but the cells may be engineered to give different coating weights on the two sides, so-called differential coatings.

The plated tin shows the matte finish at this point; it is dull and somewhat whitish. To brighten the surface appearance, the tin plate is “reflowed” by passing through a unit that heats the tin coating to just above its melting point. At the same time a thin, continuous layer of iron–tin intermetallic compound is formed between the steel and the tin layer, and corrosion properties of this intermetallic improve the subsequent performance of the tin plate.

Before being recoiled at the exit end of the process line, the tin plate undergoes two more treatments. A chromate-based passivation film is developed chemically on the surface to provide stability against excessive oxidation. The strip then receives a very thin coating of oil, which helps prevent damage to the surface by rubbing during shipment. Next the strip passes an inspection station, where optical or automatic scanning equipment notes any defects in the surface. Usually, the strip is recoiled, although sometimes it is cut into sheets at this point. More often, if cut sheets are needed, they are made as a separate operation so that the tin plate line can be operated more efficiently.

The steel strip that is the basis for tin plate is made to exacting standards. The steel impurities must be kept low to optimize the corrosion resistance and fabricability of the tin plate. The strip must also have a carefully controlled thickness, often in the range of 0.15 to 0.35 mm, and must have a consistent thickness across and along the strip. The temper of the strip must be of such value as to give the required degree of stiffness and drawability required in further fabrication.

The tin coating on the surface of the steel strip must similarly meet exacting specifications. It is typically 0.0003–0.0008 mm thick on each side, with a strong tendency in recent years toward the low end of this range. The new low-tin-coating tin plates range down to ~0.00007 mm for the tin thickness.

2. Two-Piece Can Manufacture

Can manufacture has also undergone many changes. With increased competition from alternative materials, tin plate cans must be made less expensively and, for beverage containers especially, two-piece can-making methods can be more efficient. The two-piece method uses circular blanks stamped from tin plate. These blanks are drawn with a punch and die arrangement to form a cup. The cup can then be redrawn to the final shape, or the cup walls can be ironed, which thins the tin plate sides and extends them up to form the final shape. The redrawing process produces a more uniform thickness overall and therefore is generally limited to smaller height-to-diameter ratios. The wall-ironed can has a relatively thicker bottom and relatively thinner sides and can be made into the taller, narrower profiles.

Huge presses can draw these cans from stock at high speeds, approaching 200 strokes per minute. This productivity puts great demands on the tin plate, since any defects in the stock may cause a tearing of the can side and the possible jamming of the press. Temper, drawability, and sheet thickness must all be held to close tolerances. The tin is a useful metallic lubricant in each process and also imparts more traditional benefits described below.

No matter which drawing process is used, the remaining steps are similar. The can top edge is trimmed to final dimension, a step required not only because of processing variables, but also owing to sheet anisotropy. The mechanical property difference between the direction parallel to the rolling direction and the direction transverse results in “ears,” regular height variations along the top edge.

The next step is exterior decoration and interior coating of the can body (plus bottom) “in the round.” This complex subject is discussed in Section I.A.3. As a final step, the top edge of the can may be “necked-in,” a reduction in diameter that reduces the required can end diameter, which results in materials savings and facilitates stacking of the cans in shipment, storage, and display. The cans are shipped to the filling location, where the ends (tops only) are seamed on after filling.

3. Three-Piece Can Manufacture

The three-piece method of tin can manufacture is the traditional method and still accounts for the majority of cans.

The process begins with cut tin plate sheets. Food cans often have plain exteriors, to receive paper labels after filling. When the exterior is to receive decoration, however, it is usually done on flat sheets at this time, either in-house or by custom metal decorators.

Metal-decorating processes vary greatly depending on the performance requirements for the finished can. The exterior coatings are termed inks, varnishes, lacquers, or enamels, with some overlap in definitions. There may be size coats, base coats, and so on, and the processing can be roller coating or photolithography (or both). Distortion printing can be used to allow in advance for changes in shape during container forming. All these possibilities are used to good advantage by can designers to present a package that catches the attention of the consumer and conveys the sense of the quality of the product.

The interior coating, however, is undoubtedly more crucial to product preservation. In many instances, a plain tin interior allows a controlled dissolution of tin, which protects the product from deterioration in taste, color, or other organoleptic qualities. When a tin surface is not essential, any of a variety of enamels can be applied to prevent undesired product–container interactions. Even in these cases, a tin coating underneath the enamel provides a further protective barrier in case of inadvertent damage to the enamel coating.

The coated sheets of tin plate are now slit into body blanks, which are then formed into cylinders around a mandrel. The seam is typically overlapped slightly and welded together under compression using a moving consumable intermediate copper wire electrode. Rapid progress in the technology of can welding has made it possible for can manufacturers to satisfy the demand for high production rates. The tin coating can be quite thin when it is to be welded, but it is necessary for the improved electrical contact properties it provides. Welding eliminates any concerns over solder contamination of the contents and is still gaining favor for this reason.

Soldering the side seam is the older method of manufacture and is still used on nonfood packaging and some “dry” packs. For soldering, the side seam is folded into a lock seam, fluxed and pressed flat, and wiped with a 2% tin–98% lead solder. Once again this process must be capable of high speeds. In some cases prescored blanks of two or three cans are soldered as a unit, then divided into the appropriate parts. The tin coating provides an excellent solderable coating for high-speed operation, in addition to its long-established capacity to preserve the contents.

Whether welded or soldered, the can bodies may next have the side seam sprayed with a strip of special interior enamel. This strip is cured and followed by flanging—that is, bending both top edge and bottom edge out at a right angle to allow the can ends to be seamed on as

appropriate. An overall protective interior enamel coating may be applied next and cured.

Next the can bottoms, separately produced, are put into place. These have been stamped from sheet, pressed into the needed profile, coated as required, and finally edged with a sealing compound and seamed onto the bodies. Once again the tops are seamed into place after product filling.

4. Other Products

Although the preceding description concentrates on the very important application of tin plate to food and beverage containers, a wide range of other products are packed in tin plate containers—for example, paints, aerosols, specialty chemicals, and batteries. The specific details of fabrication vary, of course, but the soldering, welding, and deep drawing capacities of tin plate are all used to full advantage by the industries involved.

Tin plate also has noncontainer uses for various corrosion-resistance applications. These include gaskets, seals, automotive air and oil filters, baking pans, graters, badges, toys, and so on. Tin plate has also been given a laminated plastic coating, which allows interesting texture effects for more decorative purposes. The deep drawing capacity of tin plate has spawned widespread use as electronic shielding against electromagnetic interference in computers, televisions, video games, and similar devices.

B. Solder

Though the second most common use of tin worldwide, solder is overtaking tin plate in the developed countries, not surprisingly in proportion to the level of electronics industry development. Although the transportation industry consumes large tonnages of solder, the higher average tin content of solder used in electronics makes the latter the more important consumer of tin.

In simplest terms, solder is a mixture of tin and lead with up to 63% tin in nearly all cases. Smaller amounts of antimony, silver, copper, cadmium, bismuth, indium, and other elements may be added for special purposes. The eutectic composition at 62% tin–38% lead melts at the lowest temperature of the tin–lead combinations, that is, at 183°C. This low temperature for metal joining has proved valuable for joining heat-sensitive electronics components with a minimal chance of heat damage.

Besides allowing joining at low temperatures, soldering also benefits from the capacity of tin to wet and alloy with a variety of useful metals of construction. Often solder is relatively inexpensive compared with other joining materials, and reasonable degrees of automation are possible. The good corrosion resistance of tin and lead in specific

environments can be used advantageously by solder alloy selection.

1. Soldering Essentials

Solder alloy selection is a complex topic. To provide a few guidelines, a nearly eutectic composition of 60 to 63% tin is standard for electronics soldering. Alloys of 50% tin are traditional for plumbing and sheet metal applications. In the 20 to 40% range, the alloys are used for general engineering purposes, and the low-tin solders, up to perhaps 10% tin, are standard for can soldering, radiator soldering, and some electronics soldering in which a two-step soldering operation is required. Pure tin has been used for can side-seam soldering of milk products and baby foods and for step soldering. Antimony or silver is generally added for higher temperature soldering or for additional strength. Adding cadmium or bismuth lowers the melting point for fusible alloys and special applications.

An important design limitation of the high-tin solders involves service temperatures in the cryogenic range. Alloys of more than ~20% tin undergo a dramatic ductile-to-brittle transition at about -100°C , so for such service temperatures low-tin solders are required.

The essential constituents of a soldered joint are the basis metals, a flux, a solder, and a source of heat. The basis metals must be clean and solderable, and if solderability of the metal is difficult (aluminum, stainless steel, case iron), it may be advisable or necessary to apply a plating of a solderable metal. Even if the metal has good solderability (copper, brass, low alloy steel), if some storage is required, a solderable coating will help preserve solderability for some time. Tin or tin-lead platings or hot-dip coatings are excellent for this purpose.

A flux is a chemical agent that removes light tarnish films on the basis metal, protects the surfaces from reoxidation during heating, and generally assists the molten solder to wet and spread over the surfaces to be soldered. Fluxes are often of proprietary compositions. For electronics they usually consist of a natural wood resin base (called "rosin base") with small amounts of halide-containing activators to improve the performance. For more difficult soldering situations and when postcleaning to remove flux residues is possible, mixtures of organic acids or of inorganic halides are used as fluxes. These more corrosive fluxes may leave corrosive residues on the soldered surfaces, and this is the reason cleaning after soldering is a necessity.

The solder composition is selected as already outlined, and for many soldering processes the solder application step is combined with the source of heating. For example, in the wave soldering process commonly used for electronics, a printed circuit board loaded with electronic

components is moved by conveyor through a machine that first sprays flux on the bottom of the board or moves the board through a wide wave of flux pumped up to meet it. Next the board passes over a wave of molten solder that supplies both heat and solder to the surface. Finally, after cooling the board passes through a cleaning section. By such a process several hundred solder joints can be made in a few seconds. Dip soldering is a rather similar process.

A large number of joints are still made by hand soldering, where the solder and heat are applied separately. The solder might be in the form of wire, stick, foil, stamped preforms (small shapes customized to the specific application), or solder paste. These forms may have the flux incorporated as a core inside or may be flux-coated or, in the case of pastes, may be mixtures of flux and solder powder. The heat can then be applied with a soldering iron, with a torch, by electrical resistance heating, with a hotplate, with an oven, or by condensation of a heated fluorocarbon (called vapor-phase soldering). There are many variations in the specific details of soldering.

A growing concern among solder users is the mechanical behavior of solder joints. Traditionally, joints to be soldered were mechanically fixed to provide support, with the solder simply a filler metal to maintain electrical continuity and corrosion resistance. In recent years, however, mechanical fixing has proved time-consuming to accomplish and, in the case of delicate components, perhaps impossible. This means that the solder joint is expected to have adequate mechanical strength. In addition, there are new joining techniques, such as surface mounting, which involves direct mounting of electronic components to a circuit board with no leads on the components. In this method the solder joint must withstand the fatigue forces caused by the differences in thermal expansion when the component materials are heated and cooled in service. In fact, in nearly all soldering applications the service conditions are becoming more demanding as designers aim for more efficiency and performance.

In transportation applications, wider use of aluminum radiators will reduce the use of the low-tin solders that have become standard. The high-lead solders used for body filling are being replaced, in part, by a high-tin (nonleaded) solder. The wider use of electronics in automobiles should increase the usage of tin-rich solders.

In construction applications high-tin (nonleaded) solders are being used more for plumbing to reduce even more the very slight chance of lead contamination of water. Some of these alloys were already in use for special high-strength requirements. Overall, the use of tin in soldering shows promise of growth in the future.

C. Other Metallurgical Applications

1. Bearings

A good bearing material is one of the best examples of an engineering compromise in modern technology. The bearing must be strong enough not to deform under fluctuating loads, yet must deform enough to conform to shaft alignment variations. The bearing must be hard enough to resist wear, but soft enough to allow dirt and contamination to become embedded in the surface rather than cause additional wear. Finally, the bearing must allow a lubricating film to be maintained, yet have high corrosion resistance against many hostile environments. Not surprisingly, bearings therefore are often two-phase structures, one a hard phase for strength and hardness and the other a soft phase for conforming and embedding. The capacity of tin to hold an oil film plus its good corrosion resistance, softness, and good alloying behavior make it a natural candidate for one phase.

Tin-based white metals (often called Babbitt alloys after an inventor of 150 years ago) generally, use 7–15% antimony plus 1–5% copper, with perhaps a few percent lead or other alloying elements. The antimony and copper combine with tin to form intermetallic compounds, which provide the hard phase for the bearing material, the remaining tin-rich phase being the soft one. These alloys are still preferred for large-equipment applications (marine diesels, steam turbine bearings, earth-moving machinery, railroad applications) in which the conformability, embeddability, and corrosion resistance of tin make it indispensable. The castability and good general fabricating properties of these alloys facilitate repair and maintenance in remote locations, another important advantage.

Tin is also combined with aluminum, at contents at 6 to 40% tin, to form a family of bearing metals. The higher tin alloys among these are used for their higher strengths (load-bearing capacities) compared with the tin-based white metals, for example, in turbines and some aircraft and combustion engine applications. The lower tin alloys are used for their improved fatigue strength, as well, in high-speed engines and pumps. A disadvantage of the latter is that they must be used against a hardened shaft to prevent undue wear.

Tin bronzes make use of the copper properties in high loadings and low-speed machinery, such as rolling mills. They can be fabricated from metal powders and can therefore be impregnated with graphite or polytetrafluoroethylene, either of which can serve as a continuous-supply lubricant, allowing (otherwise) unlubricated operation, although at relatively low loads. The bronze powder parts can also be impregnated with oil as a lubricant for sealed, low-maintenance operation.

2. Bronze

Bronze is, of course, a use for tin that predates recorded history. Even today, a significant amount of tin is used to make bronzes. Often bronze castings are used as bearings in the manner just mentioned. Other bronze castings are often used for their good corrosion resistance and wear resistance, as valves, for example, in chemical processing equipment. All these castings contain 5–12% tin, with higher tin contents wherever corrosion resistance is more important. In some applications zinc additions form alloys called gunmetals, which can be more easily cast. Still higher tin contents (12–24%) are used for statuary castings, bells, and cymbals.

Wrought bronzes with 2 to 8% tin and additions of nickel, aluminum, or magnesium are used for seawater corrosion resistance, tarnish resistance, and good electrical contact performance. Some of these alloys show age hardenability, so that they can be formed readily before aging, then hardened to achieve electrical contact springiness and wear resistance with a good electrical conductivity.

3. Tin-Based Alloys

a. Pewter. The purest tin alloy with which the typical consumer may be familiar is the modern pewter alloy. Typically, the alloy is about ~92% tin, 6–7% antimony, and 1–2% copper, sometimes with bismuth or silver additions also. Pewter is eminently castable, with gravity casting in permanent molds being the traditional method for long production runs. More recently, centrifugal casting in rubber molds has become very popular for casting figurines, jewelry, belt buckles, and other objects. Pressure die casting is another useful technique for mass production and has been used for larger forms, such as drinking cups.

Also very useful for large-scale production of pewterware are spinning and deep drawing techniques. Both make use of the high ductility of pewter, which allows stretching and bending of the metal with little work hardening. In spinning, sheets or circles of pewter are turned down on a lathe, usually onto a wooden form, the spinning metal being pressed into place manually or automatically on such lathes. Goblets, teapots, and other tableware are often spun.

Intricate shapes of cast or spun items can be created by soldering individual parts together. With practice an experienced solderer can make a joint that will be invisible once it is buffed and polished. Since the pewter alloy has no lead added, the surface will not darken but will retain its luster with only minimal care.

b. Die-casting alloys. Die casting began with tin-based alloys because of their low melting points and high

fluidity. Besides the pewter alloys just mentioned, a range of tin–antimony–lead alloys, with some copper added to certain alloys, are used for die casting. Various mechanical parts, such as gears and wheels and weights, are a few examples of tin die castings. Postage meter print wheels, in particular, make use of the capacity of tin alloys to form sharp details.

c. Fusible alloys. Combinations of several metals, especially tin, bismuth, lead, cadmium, and indium in various proportions, make a wide range of alloys, often eutectics, with melting points ranging from 47 up to 183°C. These are of obvious value as fusible links in fire alarms and sprinklers, temperature indicators, and devices to protect against overheating of equipment. Fusible alloys are also useful as seals for temperature-sensitive components and as castable tool holders and molds in metal working and plastics fabrication.

4. Tin in Ferrous Materials

Although tin is usually avoided as a steel impurity due to its embrittling effect, there are two uses for tin in ferrous materials. First, as an addition to cast iron, at about a 0.1% level, tin stabilizes pearlite. In so doing it reduces the tendency toward softening and dimensional change that may occur in high-temperature service, and, unlike some other pearlite stabilizers, it does not impair the machinability of the iron castings. These properties are made use of in automotive parts, such as engine blocks, crankshafts, and transmission components, and similar industrial equipment.

Tin powder has also proved useful as an addition with copper to iron powder metallurgy parts. The advantage of this application can be the use of a lower sintering temperature, but more often it is the closer dimensional control that tin allows. Pistons and connecting rods for refrigerator compressors are well-known applications for this process.

5. Pure Tin

A few applications of pure tin exist. Tin foil laminated on either side of lead foil is used for wine bottle capsules, the sheaths that fit over the corked end. Tin foil is also used for some electrical capacitors and for wrapping high-quality chocolates.

Some pure tin is still used for collapsible tubes for medicines and artists' paints, but the newest and most interesting application for pure tin is as a molten bath on which molten glass is cast, the "float glass" process. In this case the tin surface yields the optical flatness that eliminates the polishing formerly employed to make plate glass. The float glass process has quickly taken over the

production of such products as windows, mirrors, and automobile windshields.

D. Tin and Tin Alloy Coatings

Since tin plate, as reviewed earlier, is such a large consumer of tin, the metal tin is undoubtedly the most commonly electroplated metal. Even without tin plate, however, the uses of electrodeposited tin and tin alloys, along with other methods for applying the coatings, are important industrially. Though the tin is often introduced into the system in the form of a chemical, it ends up in metallic form in the coating so that the reasons for specifying the coating are metallurgical.

1. Tin Electrodeposits

Nearly all the applications of tin-electroplated coatings are related to the excellent corrosion resistance of the metal. Although some are purely decorative, most of these applications have a functional nature. Often the coatings are used on food equipment, where the long association of tin with food preservation gives a sense of safety.

Tin is cathodic to steel and iron and to copper, so it is usual to plate a coating that is effectively pore free. Otherwise, rusting or corrosion may occur at the plating pores. Generally, at least about an 8- μm thickness is considered necessary, but thicker coatings are often used to allow for possible abuse in service or for more severe outdoor service. Tin coatings of $\sim 8 \mu\text{m}$ on clean copper or copper alloys or on nickel alloys are used to keep these surfaces solderable for a year or more in normal storage. A fairly recent application is electrical contacts, where in certain designs tin has been substituted for the gold coatings previously used. The lubricating properties of tin are also useful, and tin has therefore become a standard coating for pistons and piston rings used in compressors and for the thread sections in oil-well piping.

Tin is usually plated from one of three electrolytes: alkaline plating solutions based on sodium or potassium stannate, acid baths based on stannous sulfate, or stannous fluoborate baths. Each bath has certain advantages and disadvantages compared with the others. Certain acid baths, for example, can deposit a bright tin plating by codepositing brightening additives with the tin. This technique bypasses the reflowing step sometimes incorporated to brighten the matte-finish deposits, but the brightening additives sometimes prove deleterious in later processing steps.

2. Tin–Lead Alloy Electrodeposits

Tin–lead electrodeposits, traditionally from fluoborate baths but more recently from some proprietary

nonfluoroborate baths as well, produce a coating that is easily soldered. This coating is, after all, essentially solder already. Tin–lead is widely used in the electronics industry for coating printed circuit boards and component leads to be soldered into those boards. It is also used by the automotive industry on radiator parts and various electrical connectors.

3. Tin–Nickel Alloy Electrodeposits

Although many tin alloy electrodeposits perform as if they were simple mixtures of the metals, the usual tin–nickel deposit is an intermetallic compound of 65% tin, deposited from a chloride–fluoride electrolyte. The coating is hard and tarnish resistant with good oil retention properties. Therefore, it has found use as an electrical contact material, for watch parts, and for precision instruments. Since it is also reasonably solderable, the plating is used in some electrical and electronic applications.

4. Tin–Zinc Alloy Electrodeposits

The tin–zinc alloy, from a stannate–cyanide bath, was originally developed as a substitute for cadmium electrodeposits. It behaves much like a mixture of tin and zinc, better than pure zinc in marine service conditions but inferior in industrial conditions. The preferred coatings are 70–85% tin.

Once again, electrical and electronic applications are common. Automotive applications, such as brake systems, make use of the resistance to hydraulic fluids.

5. Other Tin Alloy Electrodeposits

Bronze platings are important for their wear resistance and corrosion resistance in the manner outlined earlier for solid bronze. Also in the copper–tin family of alloy platings is “speculum,” at ~40% tin, which resembles silver in color and is known for its tarnish resistance. The plating conditions for speculum require close, careful control, and therefore its use is not as widespread as it might be.

Tin–cadmium electrodeposits behave rather like tin–zinc and also offer good protection in marine environments. A relatively recent development is tin–cobalt alloys, which show an excellent color match to chromium coatings.

6. Other Coating Methods

a. Immersion tin. Tin and some tin alloy coatings can be deposited by simple immersion in a chemical bath. This process utilizes a replacement of surface atoms by atoms of tin as the former pass into the bath solution. The reaction ceases when the surface has been coated with tin.

Immersion tin has been used on some electronic devices as an overlayer to whiten the appearance of tin–lead electrodeposits. It has also been used to coat the interior of narrow-diameter copper pipe where electroplating would not produce coverage. Immersion tin on steel wire functions as a bonding agent for rubber or as a lubricant for drawing processes. Immersion tin on aluminum alloys provides a base for further electroplating or on aluminum pistons helps to prevent scuffing of the surface during the early stages of use.

b. Autocatalytic tin. Considerable research has been directed toward the development of an autocatalytic method for depositing tin. Such a process would coat a surface continuing past the end point of immersion coatings to provide a thick solderable layer of tin. Improvements in the speed and stability of the present formulations could make this process a commercial reality.

c. Hot-dip coatings. The first tin coatings were probably applied by hot dipping. Some Bronze Age artifacts appear to have been tin-coated by unknown processes, but surely hot dipping is the most likely candidate. Before World War II, all tin plate was made by this process, and until recent years farmers stored their milk in hot-tin-dipped milk cans awaiting shipment to the processing plant. Institutional mixing bowls and food grinders are just two examples of items that are still given a hot-dip coating of pure tin.

Tin–lead (or solder) is applied by hot dipping also. Electronics manufacturers who have seen the quality and environmental control advantages of this coating method are expressing renewed interest in both hot tin and hot solder coatings.

The method of hot dipping is not unlike soldering in that a flux is applied to remove light tarnish films on the base metal, and the part is dipped in a pot, which heats it and acts as a supply of tin (or solder) as well. The part must be drained to give a smooth coating. Of course, the operator must possess considerable skill for the handling of large or intricate shapes, but effectively any metal that can be soldered can be hot-dipped. Obviously, the metal must be capable of withstanding molten tin temperatures. Even small parts can be handled in bulk using baskets and employing a mechanical separation step to keep the parts from sticking together.

d. Terneplate. Terneplate is a mild steel strip or sheet coated, usually by automated hot dipping, with a lead–tin alloy of 8 to 20% tin. The terneplate of lower tin content is the standard material for automobile gasoline tanks and associated parts. Terneplate of higher tin contents is used in roofing materials, signs, and other outdoor-exposure

TABLE V Properties and Preparation of Some Important Inorganic Tin Compounds

Compound	Chemical formula	mp (°C) ^a	bp (°C) ^a	Specific gravity	Typical preparation
Stannic oxide	SnO ₂	~1637	—	6.95	Thermal oxidation of tin metal
Stannic chloride	SnCl ₄	−30	114	2.23	Direct chlorination of tin metal
Stannous octoate	Sn(C ₈ H ₁₅ O ₂) ₂	—	(d)	1.26	Reaction of stannous oxide and 2-ethylhexoic acid
Sodium stannate	Na ₂ Sn(OH) ₆	—	—	3.03	Fusion of stannic oxide and sodium hydroxide plus leaching
Potassium stannate	K ₂ Sn(OH) ₆	—	—	3.30	Fusion of stannic oxide and potassium carbonate plus leaching
Stannous chloride	SnCl ₂	247	652	3.95	Direct chlorination of tin metal
Stannous sulfate	SnSO ₄	360 (d)	—	4.18	Reaction of stannous oxide and sulfuric acid
Stannous fluoride	SnF ₂	220	853	4.9–5.3	Reaction of metallic tin and hydrofluoric acid
Stannous oxide	SnO	1080 (d)	—	6.45	Reaction of stannous chloride and alkali followed by heating

^a (d), Decomposes.

items. A frequently cited example of a roofing application is the Andrew Jackson home in Tennessee, which still has the terneplate roof installed in 1835.

Even a material as old as terneplate continues to evolve. For automotive applications, terneplate producers are introducing a nickel preplate on the steel before terne coating. This preplate is said to improve the corrosion resistance and general performance of the product. For roofing applications a terne-coated stainless steel strip has become available and has been termed “nearly indestructible” in outdoor exposure tests.

E. Chemical Applications: Stabilizers and Catalysts

In recent years tin chemicals have been the most rapidly growing areas of tin usage. The consumption of inorganic tin chemicals remains larger than that of organotins, but the use of the latter is expanding rapidly. The distinction

between the two is that organotin compounds have at least one tin–carbon bond (see Tables V and VI).

1. Inorganic Tin Stabilizers

Without a stabilizer, plastics, such as polyvinyl chloride (PVC), when heated or even just exposed to light for a time, will discolor or become brittle (or both). Incorporating stannous stearate or stannous oleate into the plastic prevents heat from removing HCl from the PVC, thereby stabilizing the product. These compounds may be used in food service grades.

2. Organotin Stabilizers

A range of di- and monoalkyltin compounds are also used for stabilizing PVC, and in fact this is the largest use of any organotin compounds at present. In general, organotins containing tin–sulfur bonds have proved the most effective

TABLE VI Properties of Some Important Organotin Compounds

Compound	Chemical formula	MW	mp (°C)	bp (°C) ^a	Specific gravity	Typical applications ^b
Dimethyltin dichloride	(CH ₃) ₂ SnCl ₂	220	107	187	—	Glass coating precursor
Dibutyltin dilaurate	(C ₄ H ₉) ₂ Sn(OOCC ₁₁ H ₂₃) ₂	632	23	400 (10 mmHg)	1.05	Catalyst
Dimethyltin bis(isooctyl-mercapto-acetate)	(CH ₃) ₂ Sn(SCH ₂ CO ₂ C ₈ H ₁₇) ₂	491	—	—	1.18	PVC stabilizer
Dibutyltin bis(isooctyl-mercapto-acetate)	(C ₄ H ₉) ₂ Sn(SCH ₂ CO ₂ C ₈ H ₁₇) ₂	575	—	(d)	1.11	PVC stabilizer
Bis(tributyltin) oxide	(C ₄ H ₉) ₃ SnOSn(C ₄ H ₉) ₃	596	Less than −45	212 (10 mmHg)	1.17	Biocide, wood preservative
Triphenyltin hydroxide	(C ₆ H ₅) ₃ SnOH	367	120	—	—	Fungicide

^a (d), Decomposes.

^b PVC, Polyvinyl chloride.

heat stabilizers, and dialkyltin carboxylates are used for good light stability. These compounds, too, are considered safe enough for use in food-grade packaging materials, rigid PVC piping for potable water supply, and plastic beverage bottles.

3. Catalysts

Stannic oxide is used in the petroleum industry as a heterogeneous oxidation catalyst. Stannous octoate is widely used in the production of polyurethane foams and in room-temperature-vulcanizing silicones. Stannous oxalate and some other carboxylates catalyze other chemical reactions of industrial significance.

Among the organotins, dibutyltin dilaurate and related compounds have also been used for polyurethane foams and for silicones. Some monobutyltin compounds have proved effective as catalysts for esterification reactions.

F. Chemical Applications: Biocides

Whereas mono- and diorganotin compounds possess such low toxicities that they are perfectly acceptable for food-contact plastics stabilizers, another class of organotins, the triorganotin compounds, exhibit an interesting and useful range of toxicities to living organisms. The compounds tend to be very selective in their actions depending on the organic constituent on the tin atom. For example, trimethyltin and triethyltin compounds are most toxic to mammals, so they are not used commercially. In contrast, tributyltin and triphenyltin compounds are effective against fungi and mollusks and therefore are of use as fungicides and antifouling agents.

Triphenyltin compounds have been used as fungicides since the 1960s. Not only do they combat fungi, but they also act as antifeedants, discouraging insects from feeding on the protected plants. Other triorganotins have found use in agriculture as miticides.

Bis(tributyltin) oxide has been used since 1960 for fungal control in wood, cotton, and other cellulosic materials. Water dispersions of this compound have proved effective in preventing moss and algae from growing on stonework. New organotin compounds have been developed that are more water soluble and show promise in expanding the applications of organotins in these areas.

Marine fouling is a persistent problem for commercial and naval ships, where the drag on a ship caused by the attachment of marine organisms rapidly decreases the fuel economy. Less effective antifouling paints may increase costs by requiring frequent dry-docking and consequent loss of service. Organotin-containing paints, however, have been shown to be very effective and long-lived and well worth any increased cost in initial application. Newer

paint formulations incorporate the organotin into the paint polymer itself so that fresh antifouling agent is exposed as the paint is eroded—that is, as necessary to be effective.

A dialkyltin compound that has biocidal application is dibutyltin dilaurate, which is used as an antiwormer in poultry. It is typical of the selectivity of organotins that this treatment kills the worm parasites without harming the infected chickens or turkeys.

G. Other Chemical Applications

An important and promising application of tin is as a coating on glass. A thin layer of stannic chloride or monobutyltin trichloride is sprayed onto newly formed, still hot glass beverage bottles, jars, and glasses. The chemical is then converted to stannic oxide ($\sim 0.1 \mu\text{m}$ thick) by the heat, and this coating strengthens the glass, paying for itself in reduced breakage as the glass items are processed.

Thicker coatings ($1 \mu\text{m}$ thick or more) of stannic oxide, sometimes combined with indium oxide, combine electrical conductivity with transparency to produce deicing glass for aircraft windows or illuminated signs or, for stannic oxide alone, precision electrical resistors. The intermediate thicknesses are the most promising, however, for in this range coatings on window glass have shown excellent insulating and heat-saving capabilities to compete with sputtered coatings but at a fraction of the price. This usage is envisioned for commercial construction and also for home use, where a significant savings in home heating costs may be effected.

Tin chemicals are also used in ceramic glazes as opacifiers, in ceramic pigments, as glass-melting electrodes, as pharmaceuticals (stannous fluoride in toothpastes), and as reducing agents. There are a variety of still more minor uses in many different industries.

Interesting new applications for organotins involve the use of mono-organotins as water-repellent agents for textiles and masonry applications and as possible flame retardants and smoke suppressants in a variety of plastics and cellulosic materials. These markets offer a huge potential for an effective product.

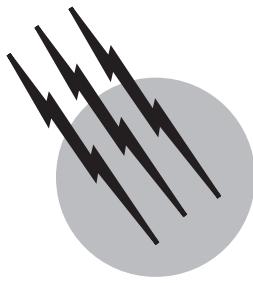
SEE ALSO THE FOLLOWING ARTICLES

CORROSION • ELECTROCHEMISTRY • MINERAL PROCESSING • MIRROR ENGINEERING

BIBLIOGRAPHY

Barry, B. T. K., and Thwaites, C. J. (1983). "Tin and Its Alloys and Compounds," Ellis Horwood, Chichester.

- Evans, C. J., and Karpel, S. (1985). "Organotin Compounds in Modern Technology," Elsevier, New York.
- Gielen, M., ed. (1990). "Tin-Based Antitumour Drugs," NATO ASI Series, Vol. 0, Springer-Verlag, Berlin.
- Harrison, P. G., ed. (1989). "Chemistry of Tin," Blackie and Son, Glasgow.
- ITRI. (1983). "Guide to Tinplate," International Tin Research Institute, London.
- Lehmann, B. (1990). "Metallurgy of Tin," Lecture Notes in Earth Sciences Series, Vol. 32, Springer-Verlag, Berlin.
- Leidheiser, H. (1979). "The Corrosion of Copper, Tin and Their Alloys," reprint, Krieger, Huntington, New York.
- Price, J. W. (1983). "Tin and Tin-Alloy Plating," Electrochemical Publications, Ayr, Scotland.
- Robertson, W. (1982). "Tin: Its Production and Marketing," Croom Helm, London.
- Schumann, H., and Schumann, I. (1988). "Gmelin Handbook of Inorganic and Organometallic Chemistry," 8th ed., Gmelin Handbook Series, Vol. 16, Springer-Verlag, Berlin.
- Thwaites, C. J. (1977). "Soft-Soldering Handbook," International Tin Research Institute, London.
- Wright, P. A. (1983). "Extractive Metallurgy of Tin," 2nd ed., Elsevier, New York.



Uranium

William L. Chenoweth

Nuclear Minerals Consultant

Thomas C. Pool

International Nuclear, Inc.

- I. Physical Properties
- II. Utilization and Production
- III. Geologic Environment
- IV. Exploration and Refining

GLOSSARY

- Alaskite** A commonly used term for a granitic rock containing only a few percent of dark minerals.
- Aquifer** A body of rock that is sufficiently permeable to conduct ground water and to yield significant quantities of water to wells.
- Breccia** A coarse-grained rock composed of angular broken rock fragments.
- Carbonatite** A carbonate rock of magmatic origin.
- Carnotite** A canary-yellow to greenish-yellow radioactive mineral-potassium, uranium, vanadate—commonly found in sandstone. An ore of uranium, vanadium, and a source of radium.
- Collapse-breccia pipe** A circular, pipelike structure containing a breccia of sedimentary rocks which have collapsed downward into a cave or solution cavity.
- Magna** Naturally occurring mobile rock material generated within the earth.
- Magmatic segregation** Concentration of uranium in certain parts of a magna during its cooling and crystallization.
- Pegmatite** An exceptionally coarse-grained igneous rock, with interlocking crystals.

Phosphate rock Any rock that contains one or more phosphatic minerals of sufficient purity and quantity to permit its commercial use as a source of phosphatic compounds.

Pitchblende A noncrystalline variety of the mineral uraninite, a uranium oxide. Commonly occurs in vein deposits. An ore of uranium and a source of radium.

Placer A surficial mineral deposit formed by mechanical concentration of mineral particles.

Proterozoic A division of geologic time consisting of the younger Precambrian. Dated as occurring between 2500 and 570 million years ago.

Radium A radioactive, metallic element found in very small amounts in certain uranium ores. Atomic number 88, atomic weight 226.05.

Radon A radioactive gas formed by the disintegration of radium. Atomic number 86, atomic weight 222.

Solution mining The recovery, by chemical leaching, of the components of a mineral deposit without physical extraction of the rock. Also referred to as *in situ* mining.

Tailings The powdered rock and other waste products from a milling process.

Tuff A consolidated rock composed of volcanic ash and other pyroclastic material.

Unconformity A substantial break or gap in the geologic record where a rock unit is overlain by another that is not next in the stratigraphic succession.

Uranium hexafluoride A compound (UF_6) obtained by chemical treatment of uranium oxide (U_3O_8), which forms a vapor at temperatures above 56°C.

Uranium ore Rock containing uranium mineralization that can be mined economically.

Vanadium A metallic element used to increase the strength of steel. A constituent of the mineral carnotite. Atomic number 23, atomic weight 50.95.

Vein The mineral filling of a cavity such as a fault or other fracture in a host rock. Commonly tabular or sheetlike in form.

Yellowcake The common name for uranium concentrate produced at uranium mills. A yellow or brown powder containing 60–75% uranium oxide (U_3O_8). A chemical precipitate, commonly a ammonium diuranate or magnesium uranite.

URANIUM is a radioactive, metallic element that is the basic raw material for atomic energy. Uranium ores were originally sought as a source of radium for medical applications. In the 1940s and 1950s, the major use of uranium was for atomic weapons. Today most uranium is used as a fuel for nuclear power plants. The heat produced by the nuclear reactor in the plant is used to generate steam, which is then used to generate electricity. One pound of natural uranium can produce as much energy as about 14,000 pound of coal.

Total world uranium supply for the period 1947 through 1999 amounts to 5.1 billion pounds U_3O_8 . Historically, the United States has been the world's largest uranium producer with a total production through 1999 in excess of 900 million pounds U_3O_8 . Canada ranks second historically, but since 1984 has ranked number one annually and in 1999 produced over 20 million pounds U_3O_8 . Other important current producers are Australia, Namibia, Uzbekistan, Kazakhstan, Niger, and Russia. United States production ranks eighth at just under 5 million pounds U_3O_8 per year. Production of uranium exceeded consumption through 1990 and large inventories were accumulated by governments, producers, and consumers. Consumption now exceeds production and inventories are being reduced.

I. PHYSICAL PROPERTIES

Uranium is a heavy, silvery-white, radioactive, metal that consists of three semistable isotopes U-238, U-235, and

U-234. It is an important energy source because fission of isotope U-235 releases large amounts of energy. This readily fissionable nuclide constitutes only about 0.7% of natural uranium. The isotope U-238 makes up most of the remaining 99.3% and the third, U-234, only about 0.005%. U-238 is not readily fissionable, but under neutron bombardment converts to plutonium-239, which is fissionable.

Uranium has been known since 1789, when it was discovered by a German chemist named Martin Heinrich Klaproth in pitchblende from a mine in Germany. The first pure uranium was isolated in 1984 by the French chemist, Eugene Melchior Peligot. He named it for the planet Uranus. In 1896, a French scientist, Antoine Henri Becquerel, discovered its property now known as radioactivity. Uranium has an atomic weight of 238.07, atomic number 92, and the symbol U.

II. UTILIZATION AND PRODUCTION

A. Uses

The element radium, a daughter of uranium decay, was discovered by Marie and Pierre Curie in 1898 in pitchblende from Czechoslovakia, where the mineral had been known since 1727. In the early 1900s radium became important in medical therapy. This led to the search for uranium ore as a source for radium. The first important sources of radium outside Czechoslovakia were the carnotite deposits on the Colorado Plateau in western Colorado and eastern Utah from which about 67,000 tons of ore was produced during 1898–1923. This ore yielded about 202 g of radium.

One gram of radium could be recovered from 200 to 300 tons of carnotite ore averaging 2.00% uranium oxide (U_3O_8).

In 1923, the United States deposits were supplanted as the source of radium by the large and rich Shinkolobwe vein deposit in the Belgian Congo (now Congo). In 1933, production began from another vein deposit, the Eldorado at Port Radium, Northwest Territories, Canada, and thereafter the market was shared by Canada and the Belgian Congo. Only minor amounts of domestic carnotite ore were mined from 1924 to 1935.

The mining of Colorado Plateau carnotite ores increased markedly in 1936, owing to the demand for vanadium by the steel industry. Prior to 1942, the uranium was used chiefly for coloring glass and ceramic glazes. An ample supply was obtained by recovering some uranium from ores mined at first for radium and later for vanadium.

In 1942, controlled nuclear fission was demonstrated, and uranium had two new and vastly more important uses: as an explosive by the military and as a source of heat to produce steam for generating electricity. In order to obtain uranium for developing atomic weapons, the United States

Army Corps of Engineers, in 1943, began recovering uranium from the tailings discarded from radium and vanadium processing. Uranium was also obtained from the Eldorado mine in Canada and the Shinkolobwe mine in Congo.

The United States Atomic Energy Commission (AEC) was created in 1947, and as part of their activities, began a large program to acquire uranium for the weapons program. The price schedules, bonuses, and other incentives offered by the AEC created a prospecting boom in the 1950s unsurpassed in the mineral industry. Hundreds of deposits were discovered in the western United States. Uranium consumption for nuclear fuel increased rapidly during the 1970s, 1980s, and 1990s, but has recently leveled off at about 160 million pounds U₃O₈ per year. No major increases are forecast for the future as the installation of new nuclear plants is expected to be balanced by decommissioning of older plants as well as by shutdowns due to deregulation of the electricity industry and antinuclear activism.

B. Production

In the early 1960s the transition from weapons to nuclear power began and in 1970 the AEC terminated its procurement program. From 1947 to 1970 the AEC purchased about 349 million pounds U₃O₈ from domestic producers. The principal producing states were New Mexico, Utah, Wyoming, Colorado, and Arizona. An additional 235 million pounds as purchased from foreign producers, mainly Canada and the Republic of South Africa. Beginning in 1966, the AEC allowed producers to sell uranium to electric utilities. Since 1971 all domestic uranium production has been for the nuclear power generating industry.

Uranium needed for anticipated nuclear power plants created a worldwide prospecting effort beginning in the mid-1960s extending to the early 1980s. This effort was highly successful in Australia, Canada, Central Africa, and the United States.

As new deposits were located and developed, uranium production increased gradually until 1975 as military needs were phased out and nuclear power programs began to develop. From 1975 to 1980 uranium production throughout the world increased dramatically and by 1980 production reached an all time high of 165 million pounds U₃O₈, of which the United States produced 43.7 million pounds. By 1980 it became clear that the anticipated growth of nuclear power was not going to materialize. A slow down in economic growth, energy conservation, and concern for the environment creating opposition to nuclear power were all responsible for the change in growth estimates. As a result production had exceeded demand, and

uranium concentrate was stockpiled. Since 1980, production has declined especially in the United States, where domestic producers could not compete with lower cost foreign producers such as Canada and Australia. In addition, the collapse of the Russia resulted in the sale of large quantities of stockpiled Soviet uranium to western buyers. These sales aggravated the existing oversupply situation and prices declined to record constant-dollar lows by the early 1990s. Both prices and production have shown some minor increases in the late 1990s. Increasing competition for uranium producers is provided by uranium from secondary sources such as reprocessing and, particularly, from the conversion of nuclear weapons to nuclear fuel. A historic pact between the United States and Russia in 1994 provided for the conversion of at most one-half of Russia's nuclear weapons into nuclear fuel and the purchase of most of that nuclear fuel by the United States. This increases the net supply of uranium by almost 25% from 1995 to 2010.

Canada is the leading uranium producer and exporter in the world. The United States-Canada Free Trade Agreement, ratified in December 1988, has allowed Canadian uranium to be imported into the United States without any restrictions, much to the regret of domestic producers. Canadian production is now exclusively from high-grade unconformity deposits in Saskatchewan. Uranium production in Australia was hindered from 1983 to 1996 by the Australian Labor Party's (ALP) antinuclear policy which allowed only three specified uranium mines to operate. The ALP's defeat in a 1996 election opened the way for additional production. South Africa previously marked third in uranium production, but has been declining due to lack of markets created by political sanctions. Uranium production in Soviet Bloc countries remained "Top Secret" through the late 1980s until the reunification of Germany, "glasnost," "perestroika," and, ultimately, the collapse of the Soviet Union provided an increasing flow of information. East Germany, Czechoslovakia, and the central Asian republics of Kazakhstan and Uzbekistan produced most of the Soviet uranium supply which totals about 1.8 billion pounds U₃O₈ through 1999. Annual Soviet Bloc production exceeded 50 million pounds of U₃O₈ per year in the late 1980s but declined rapidly thereafter as neither the Soviet system of production nor the relatively low grade deposits being exploited could compete in a market economy. Remaining production includes underground mining of vein-stock work deposits in Siberia and solution mining of sandstone deposits in Kazakhstan and Uzbekistan.

World uranium production in 2000 is expected to be about 91 million pounds U₃O₈ ([Table I](#)) while consumption is estimated at 162 million pounds. Uranium from military and civilian inventories will make up the difference. Besides the use in nuclear weapons and as a fuel for

TABLE I Estimated World Uranium Production, 2000

Country	U_3O_8 (millions of pounds)	Type(s) of deposit(s)
Africa ^a	17.8	Disseminated magmatic, sandstone, quartz-pebble conglomerate
Australia	20.4	Proterozoic unconformity, Proterozoic stratabound, sandstone
Canada	26.0	Proterozoic unconformity
Former Soviet Union ^b	16.2	Disseminated magmatic, vein, sandstone, phosphate
Other countries ^c	6.4	Sandstone, vein, disseminated magmatic, phosphate
United States	4.1	Sandstone, vein
Total	90.8	

^a Namibia, Niger, and South Africa.

^b Kazakhstan, Russia, Ukraine, and Uzbekistan.

^c Argentina, Belgium, Brazil, China, Czech Republic, France, India, Pakistan, Portugal, Romania, and Spain.

nuclear power plants, uranium is also used to produce various radioactive isotopes for medical and scientific research.

III. GEOLOGIC ENVIRONMENT

A. Geochemical Cycle

The uranium content of the crust of the earth is about 2 parts per million (ppm). Uranium originates in magmas where it is mostly in the tetravalent state. As magma crystallizes, the large size of the tetravalent uranium ion prevents it from entering the crystal lattices of common rock-forming minerals except as minor occlusions; instead, part of it is deposited in an intergranular rim on rock-forming minerals, and accessory minerals, or forms its own minerals. The remaining uranium is concentrated in late magmatic differentiates, where some uranium forms its own minerals in pegmatites and veins. In veins, iron is commonly associated with uranium, and in some places copper, lead, zinc, molybdenum, and cobalt accompany uranium. Uranium can also enter the cycle through volcanic eruptions. The resulting ash beds, or tuffs, can contain appreciable uranium.

The accessory uranium-bearing minerals of igneous rocks generally resist oxidation, and weather out to be washed into detrital sediments, or under special circumstances, to become concentrated in placers, some of which contain appreciable uranium. On the other hand, the uranium-bearing films that form on minerals in igneous rocks, including tuffs, and the uranium minerals in peg-

matites and veins oxidize readily, and water-soluble hexavalent uranium is released to surface and groundwaters. As these waters circulate through sediment some uranium may be absorbed by clay minerals and carbonaceous matter, and some is precipitated chemically or by evaporation. Reduction of hexavalent uranium to the tetravalent state and its subsequent precipitation from groundwater is the mechanism by which the uranium deposits in sandstones are thought to have been formed. Iron is universally precipitated with uranium in this part of its geochemical cycle, and in some places copper, molybdenum, selenium, and chromium are similarly precipitated.

Uranium-bearing waters also escape to the ocean where uranium may be precipitated with phosphatic sediments, taken up by organisms, or absorbed by carbonaceous mud. In the virtually oxygen-free atmosphere of early Precambrian time, it is postulated that uraninite from pegmatites and veins did not oxidize, but rather formed placer concentrations, commonly with gold and thorium-bearing minerals.

B. Mineralogy

Uranium in unoxidized black ores is tetravalent and in most deposits it occurs as uraninite (an oxide of uranium) and coffinite (a uranium silicate). Pitchblende is a massive, noncrystalline variety of uraninite. The multiple oxides brannerite (oxide of uranium, titanium, thorium, rare earths, and other elements) and davidite (oxide of titanium, iron, and uranium) are the main uranium minerals in a few unoxidized ores. Tetravalent uranium substitutes for thorium and other elements in minerals such as monazite, uranothorite, multiple oxides of niobium, and tantalum, and for calcium in carbonate fluorapatite. It also occurs in unidentified organic compounds in many coaly rocks and marine black shales.

Under oxidizing conditions tetravalent uranium changes to hexavalent uranium and forms oxide, vanadate, arsenate, silicate, sulfate, and carbonate compounds, most of which are hydrous and many of which are bright yellow or green. Carnotite (potassium uranium vanadate) and tyuyamunite (calcium uranium vanadate) are the most common and abundant. Other common minerals include uranophane (a hydrous silicate) and metazeunerite (an arsenate).

C. Types of Deposits

Uranium has been found in a variety of geologic settings throughout the world. The major resources occur in seven categories of ore types: quartz-pebble conglomerate deposits, Proterozoic unconformity-related deposits, disseminated magmatic deposits, vein deposits, sandstone deposits, surficial deposits, and other types of deposits.

The quartz-pebble conglomerate ores occur in basal Proterozoic beds which occur unconformably above the Archean basement rocks. The uranium minerals in these deposits are believed to be of placer origin. This type of deposit was formerly mined in the Elliot Lake area of Ontario, Canada, and is being mined in the Witwatersrand basin of South Africa. In the latter region, the uranium is recovered as a by-product of gold mining. Gold occurs with the uranium in the conglomerates.

Deposits of the Proterozoic unconformity type occur in metamorphic rocks below major erosional unconformities in Proterozoic rocks. These unconformities were developed during a worldwide period of deformation, about 1800–1600 million years ago. Deposits of this type contain the huge, high-grade resources in the Athabasca basin in northern Saskatchewan, and in the Alligator Rivers region of the Northern Territory, Australia.

The deposits included in the disseminated magmatic type are those associated with granites, syenites, pegmatites, carbonatites, and volcanic rocks. The largest deposit of this type is Rossing in Namibia where uranium occurs in a pegmatitic granite and alaskite. In Brazil, uranium is produced from alkaline igneous rocks at Pocos de Caldas.

Vein deposits of uranium are those in which uranium minerals fill cavities such as faults, fissures, breccias, and stock works. The massive veins of pitchblende were historically mined in Czechoslovakia, Congo, and at Port Radium in the Northwest Territories, Canada. Vein deposits are important in France, Portugal, Spain, and elsewhere in Europe.

The collapse-breccia pipe deposits have been mined in northwestern Arizona and are classed as vein deposits. They occur in sedimentary rocks of Paleozoic age.

Most of the ore deposits of the sandstone type are contained in rocks that were deposited in fluvial or marginal marine conditions. Lacustrine and eolian sandstones contain some deposits but they are not abundant. The host rocks are almost always medium to coarse-grained, poorly sorted sandstones containing pyrite and organic matter of plant origin. The host sediments are commonly associated with tuffs.

Sandstones of Triassic and Jurassic age in Colorado, Utah, Arizona, and New Mexico, and of Tertiary age in Wyoming and south Texas account for most of the past uranium production in the United States. The vanadium content of certain Jurassic sandstones (Salt Wash Member of the Morrison Formation) in southwestern Colorado and southeastern Utah makes these ore deposits important for both uranium and vanadium. Other important sandstone deposits occur in Cretaceous and Permian rocks in Argentina, in Carboniferous rocks in Niger, in sandstones of Precambrian age in Gabon in Tertiary sediments in Japan, and in Permian age rocks in Europe. Ma-

jor resources in sandstone are in the United States and Niger.

Uraniferous surficial deposits may be broadly defined as uraniferous sediments, usually of Tertiary or Recent age, which have not been subjected to deep burial and may or may not have been calcified to some degree. Uranium deposits of this type, associated with calcrete, occur in Western Australia and Namibia. The young, organic sediment deposits associated with peat bogs in North America and Scandinavia are included in this type. At the present time, no deposits of this type are being mined.

Other types of deposits are those that cannot readily be classified into the types already described. These include uranium deposits which occur in limestone and limestone karst terrain as phosphatized fractions of limestone. Uranium also occurs in marine phosphorite, carbonaceous shale, and lignite.

The huge Olympic Dam copper-uranium-gold deposit in South Australia is unique and is placed in this class. This deposit occurs in a thick sequence of unmetamorphosed sedimentary breccias of Proterozoic age. Until it is studied further it is tentatively called a Proterozoic stratabound deposit.

IV. EXPLORATION AND REFINING

A. Exploration

Economic concentrations of uranium on the surface are easily detected using radiation detection equipment in vehicles, airplanes, and carried on foot. Buried deposits are explored for using bore holes based on geologic evidence and/or theory. Deposits at depths in excess of 3000 ft are being sought. A principal technique in uranium exploration involves the measurement of radioactivity in holes drilled to evaluate a prospective target. Systematic logging of boreholes with a variety of geophysical tools, including gamma ray, self-potential, resistivity, electromagnetic, and other surveys, is a standard practice in uranium exploration. Modern exploration procedures also include detailed geological mapping, geochemical surveys, analysis of borehole cuttings and cores in the field and laboratory, and economic and engineering studies making use of pertinent data developed in the search for new uranium deposits.

B. Mining

Uranium ore is mined using open pit and underground methods similar to those used for mining other minerals and coal. *In situ* leaching of permeable sandstone deposits accounts for an increasing share of production. The depth, grade, and size of the deposit determines the type of mining to be used.

Depths of open pit uranium mines range from a few feet to depths in excess of 300 ft. Open pit mining is currently in use today in northern Australia, Niger, Namibia, Saskatchewan Canada, Spain, and Argentina.

Access to deeper ore deposits is by vertical shafts or by subhorizontal declines. Vertical shafts are used in Australia, France, Canada, Russia, and South Africa. Vertical shafts in South African gold mines where uranium is produced as a by-product of gold may exceed 7000 ft in depth.

Declines sloping 15–25° are used in Niger and the United States for access to relatively shallow underground mines. These declines require cheaper hoisting installations than vertical shafts, and for that reason find favor among small-scale operators, especially on the Colorado Plateau region of Colorado and Utah.

Horizontal access by adit is common where ore bodies are near or crop out of canyon rims. Where feasible, such access provides the simplest and cheapest method of reaching and extracting ore in underground mines. This method of mining is used throughout the Colorado Plateau especially in Colorado and Utah. Radon gas, emitted by uranium ores, is a hazard to miners and must be monitored and controlled by circulating large amounts of fresh air in underground mines.

Uranium is recovered from the groundwater that enters deep underground mines such as those in New Mexico. The water is collected in sumps and pumped to the mill for processing. Heap leaching is also used at some mines to recover uranium from lower grade ores. In this procedure, acid solutions are allowed to percolate through a mound of crushed ore. The uranium-bearing solutions are collected in a sump at the bottom of the mound and sent to the mill.

Uranium is also produced by solution or *in situ* mining methods. This process involves circulating weakly alkaline or acid solutions through an ore body continued within a permeable sandstone aquifer. The uranium in the rock is dissolved by the chemicals and the uranium-bearing solution is pumped to the surface from a production well. On the surface the uranium is precipitated as yellowcake. A “five spot pattern” of wells is commonly used. The injection wells are at the corners of a square pattern with side dimensions ranging from 50 to 200 ft. The production well is in the center of the square. Wells are 6–8 in. in diameter and are lined with polyvinyl chloride or fiberglass-reinforced pipe. A series of wells are grouped into wellfield units. Several wellfield units are usually required to mine a deposit. Uranium is recovered from solution by ion exchange in a small surface plant and precipitated, dried, and packaged into steel drums. This technique results in little surface disturbance and only small amounts of waste are generated.

After the uranium has been extracted, additional solutions are injected into the host rock to flush and restore the aquifer to its original geochemical state. In the United States, solution mining, or *in situ* leaching, has been used successfully in south Texas, Wyoming, and Nebraska at depths up to 900 ft. Uranium production in the central Asian republics of Kazakhstan and Uzbekistan is now solely by solution mining. Solution mining accounted for 18% of world uranium production and 78% of United States production in 1999.

C. Processing

Uranium-bearing ore from open pit and underground mines is delivered to nearby mills by truck or rail for processing. This involves crushing and grinding the rock to a fine size and leaching it with various chemicals to extract the uranium. The uranium is precipitated as a concentrate known as yellowcake. Other elements such as vanadium, molybdenum, gold, and the rare earths are recovered at some uranium mills, depending on the nature of the ore. Following the leaching process, the finely ground, barren rock, now called tailings, is deposited either into specially designed containment ponds or back into the mines. In the United States, considerable uranium, 45 million pounds U_3O_8 , has been recovered as a by-product of phosphoric acid produced in Florida and Louisiana from phosphate rock mined in Florida. During 1999, however, all remaining uranium from phosphoric acid facilities was demolished, as production costs were assessed to be noncompetitive in the current market. Uranium has also been produced as a by-product of: copper in the United States, South Africa, Australia, and India; gold in South Africa; processing of Moroccan phosphate in Belgium; and phosphate and rare earths in Kazakhstan. Molybdenum is frequently recovered as a by-product of uranium.

D. Nuclear Fuel

Most nuclear fuel is produced from uranium by a series of processes including: conversion, enrichment, and fabrication. Yellowcake from the conventional mills, solution mining, and by-product operations is shipped to conversion facilities where it is converted to uranium hexafluoride (UF_6).

Uranium hexafluoride is a solid at room temperature but forms a gas when heated. In gaseous form, the concentration of the fissionable isotope ^{235}U , can be increased from the natural level of 0.711% to nuclear fuel levels of 3.0–5.0% by either a diffusion or centrifuge process. In the United States by process of diffusion, gaseous UF_6 is passed through a series—or cascade—of porous membrane filters. Because UF_6 molecules containing the U-235 isotope diffuse through the filters more readily than those

containing the U-238 isotope, the diffusion process eventually results in two product streams of UF₆. Compared to the original feed material, one product stream is relatively enriched in the isotope U-235, and the other is relatively depleted in U-235.

Enrichment of the ²³⁵U isotope is necessary because the amount of fissile U-235 in natural uranium is too low to sustain a nuclear chain reaction in light-water reactors. By contrast the ²³⁵U content of nuclear weapons is typically in excess of 90%.

At the fuel fabrication plant, the enriched UF₆ is converted to uranium dioxide (UO₂). The uranium dioxide is compressed into solid, cylinder-shaped pellets, which are placed in hollow rods made of a zirconium stainless steel alloy. These rods, which are grouped into fuel rod assemblies, are shipped to nuclear power plants for use as nuclear reactor fuel. One pound of natural uranium can produce as much energy as about 14,000 lb of coal.

SEE ALSO THE FOLLOWING ARTICLES

MINERAL PROCESSING • MINERALOGY AND INSTRUMENTATION • MINING ENGINEERING • NUCLEAR FUEL CYCLES • NUCLEAR POWER REACTORS

BIBLIOGRAPHY

Energy Information Administration (1999). "Uranium Industry Annual, 1998," U.S. Dept. Energy Rep., DOE/EIA-0478(98). U.S. Govt. Printing Office, Washington, D.C.

- Ferguson, J., ed. (1984). Proterozoic unconformity, and stratabound uranium deposits, Int. Atomic Energy Agency Rep. TECDOC-315.
- Finch, W. I., and Davis J. F., eds. (1985). Geological environments of sandstone-type uranium deposits, Int. Atomic Energy Agency Rep. TECDOC-328.
- Fuchs, H. D., ed. (1986). Vein-type uranium deposits, Int. Atomic Energy Agency Rep. TECDOC-361.
- International Atomic Energy Agency (1988). INTURGEO: The international uranium geology information system, Int. Atomic Energy Agency Rep. TECDOC-471.
- International Atomic Energy Agency (1989). *In situ* leaching of uranium: Technical, environmental and economic aspects, Int. Atomic Energy Agency Rep. TECDOC-492.
- International Atomic Energy Agency (1989). Uranium resources and geology of North America, Int. Atomic Energy Agency Rep. TECDOC-500.
- Merritt, R. C. (1971). "The Extractive Metallurgy of Uranium," Colorado School of Mines Research Institute, Golden, Colorado.
- OECD Nuclear Energy Agency and the International Atomic Energy Agency (1983). "Uranium Extraction Technology," Organization for Economic Co-Operation and Development, Paris.
- OECD Nuclear Energy Agency and the International Atomic Energy Agency (1997). "Uranium Resources, Production and Demand," Organization for Economic Co-operation and Development, Paris.
- Pool, T. C. (2000). "Uranium, No Relief In Sight," *Eng. Mining J.*, in press.
- Pretorius, D., ed. (1987). Uranium deposits in Proterozoic quartz-pebble conglomerates, Int. Atomic Energy Agency Rep. TECDOC-427.
- Toens, D., ed. (1984). Surficial uranium deposits, Int. Atomic Energy Agency Rep. TECDOC-322.
- Vance, R. (1997). "Canada's Uranium Industry," Uranium and Radioactive Waste Division, Energy, Resources Branch, Natural Resources, Canada, Ottawa, Canada.