PRINCIPLES

OF GEO

CHEMISTRY

FOURTH
EDITION

MASON

MOORE

PRINCIPLES

MASON MOORE

FOURTH
EDITION

OF GEO

CHEMISTRY

PRINCIPLES OF GEOCHEMISTRY

FOURTH EDITION

BRIAN MASON

Curator, Department of Mineral Sciences
Smithsonian Institution

CARLETON B. MOORE

Department of Chemistry
Department of Geology
Arizona State University

175 YEARS OF
1807 1982
PUBLISHING

PREFACE

This book is designed to be used as an introduction to geochemistry by undergraduate and graduate chemistry and geology students as well as by those from other sciences. It summarizes the significant facts and ideas concerning the chemistry of the earth and synthesizes these data into an account of the physical and chemical evolution of the earth. The aim of the book is to explore an interesting, complex, and changing subject. An attempt has been made to cover the entire field of geochemistry. However, not all subjects have been covered with equal detail because to properly cover the facts and theories would require a volume far too long to be suitably used as a text. Some areas of geochemical interest such as phase diagrams and ore deposits have been adopted by other courses in the geology curriculum, and thermodynamics, a key to understanding geochemical processes, requires the depth of presentation given in physical chemistry courses.

The book begins with a chapter describing the scope and subject matter of geochemistry and giving a brief account of its development. The next chapter deals with the earth as a planet and its relationship to the solar system and the universe as a whole and is followed by one discussing the internal structure of the earth and its composition. From the data presented in these two chapters an account is given of the relative abundances of the elements and isotopes, in the earth and in the universe as a whole, and on the basis of the evidence an attempt is made to present a logical account of the probable pregeological history of the earth. The aim of these first chapters is to provide a background for the remainder of the book, which is concerned with materials and processes at and near the surface of the earth. Because geochemistry is to a large extent the application of physicochemical principles to processes on and within the earth, some account of these principles is given in the next chapter, with special reference to the chemistry of the solid state. The geochemistry of igneous rocks is then treated, followed by a chapter on sedimentation and sedimentary rocks. A chapter new to the fourth edition, on isotope geochemistry, includes discussions of stable isotopes and geochronology. Subsequent chapters deal in turn with the atmosphere, hydrosphere, and biosphere. All of these are inseparably intertwined in terms of their and the earth's evolution. There is no correct order in which to present some of these data. Like the geochemical cycle itself, some material

V

must be returned to at least a second time to properly relate it to other systems. The geochemistry of the hydrosphere is essentially the geochemistry of seawater and poses such fundamental questions as the mode of origin and evolution of the ocean. The geochemistry of the atmosphere is concerned with the nature of the primeval atmosphere and the changes it has undergone through interaction with the hydrosphere, the biosphere, and the lithosphere. The geochemistry of the biosphere involves a discussion of the amount of organic matter, its composition, and the role of organisms in the concentration and deposition of individual elements. Then follows a chapter on metamorphism and metamorphic rocks, and the final chapter is a brief summary and synthesis in terms of the geochemical cycle including environmental geochemistry.

The emphasis throughout is on interpretation rather than on description, on what is yet to be learned as well as what is already known. In the years since the previous edition was published, gigantic strides in the earth and space sciences have taken place. The concept of global tectonics has been developed and widely accepted, other planetary bodies have been visited, sampled, and analyzed, and the ocean bottoms have been sampled. The principles of geochemistry remain the same even though the subject of geochemistry is constantly changing.

We are aware of many gaps in our geochemical knowledge. These problems can be solved only by the combined efforts of chemists and physicists as well as by geologists. It is hoped that this text will guide them to some of the problems.

This book is necessarily a compilation from many sources, sources that include not only published books and papers but also correspondence and conversations with many individuals. Reviewers of earlier editions provided a direct contribution to this edition. A particularly valuable and smoothing effect was provided by Professor Henry Faul who read the manuscript for this fourth edition.

Teachers of geochemistry undoubtedly have developed their own approaches to the subject. We trust that *Principles of Geochemistry* will provide a useful framework upon which they may develop their chosen topics and apply their practical teaching techniques so that the field of geochemistry may continue to grow and remain vigorous.

Brian Mason
Carleton B. Moore

CONTENTS

1 INTRODUCTION

THE SUBJECT OF GEOCHEMISTRY

In the simplest terms geochemistry may be defined as the science concerned with the chemistry of the earth as a whole and of its component parts. At one and the same time it is both more restricted and also more extensive in scope than geology. Geochemistry deals with the distribution and migration of the chemical elements within the earth in space and in time. The science of the occurrence and distribution of the elements in the universe as a whole is called cosmochemistry. To properly understand the chemistry of the earth, it is important that we know as much as possible about the chemistry and history of the sun and the other planetary bodies in the solar system as well as the chemistry of the stars and interplanetary and interstellar space.

Clarke, in *The Data of Geochemistry,* defined the subject in a more restricted form:

> Each rock may be regarded, for present purposes, as a chemical system in which, by various agencies, chemical changes can be brought about. Every such change implies a disturbance of equilibrium, with the ultimate formation of a new system, which, under the new conditions, is itself stable in turn. The study of these changes is the province of geochemistry. To determine what changes are possible, how and when they occur, to observe the phenomena which attend them, and to note their final results are the functions of the geochemist. . . . From a geological point of view the solid crust of the earth is the main object of study; and the reactions which take place in it may be conveniently classified under three heads—first, reactions between the essential constituents of the earth itself; second, reactions due to its aqueous envelope; and third, reactions produced by the agency of the atmosphere.

1

V. M. Goldschmidt described geochemistry in the following terms:

> The primary purpose of geochemistry is on the one hand to determine quanti-
> tatively the composition of the earth and its parts, and on the other to discover
> the laws which control the distribution of the individual elements. To solve these
> problems the geochemist requires a comprehensive collection of analytical data
> on terrestrial material, such as rocks, waters, and the atmosphere; he also uses
> analyses of meteorites, astrophysical data on the composition of other cosmic
> bodies, and geophysical data on the nature of the earth's interior. Much valu-
> able information has also been derived from the laboratory synthesis of minerals
> and the investigation of their mode of formation and their stability conditions.

The main tasks of geochemistry may be summarized thus:

1. The determination of the relative and absolute abundances of the elements and
    of the atomic species (isotopes) in the earth.
2. The study of the distribution and migration of the individual elements in the
    various parts of the earth (the atmosphere, hydrosphere, crust, etc.), and in min-
    erals and rocks, with the object of discovering principles governing this distri-
    bution and migration.

To some degree the scope of geochemistry has been defined by the historical
development of geology. Because geochemistry as a formal discipline was relatively
late in developing, related subject areas, especially mineralogy and petrology, often
cover subject matter which might well fall within the area of geochemistry. Empir-
ical geochemical studies of element distribution on the earth's surface have played
a significant role in mineral exploration.

THE HISTORY OF GEOCHEMISTRY

The science of geochemistry has largely developed during the present century; never-
theless, the concept of an autonomous discipline dealing with the chemistry of the
earth is an old one, and the term "geochemistry" was introduced by the Swiss chem-
ist Schönbein (discoverer of ozone) in 1838. The history of geochemistry naturally
includes much of the history of chemistry and geology. Because geochemistry is
basically concerned with the chemical elements, their discovery and recognition are
landmarks in the history of the subject. The modern concept of an element can be
said to date from Lavoisier's definition in his "Traité elementaire de Chimie" (1789),
although some seventeenth- and eighteenth-century scientists certainly understood
the distinction between elements and compounds. Lavoisier recognized the following
31 elements: O, N, H, S, P, C, Cl, F, B, Sb, Ag, As, Bi, Co, Cu, Sn, Fe, Mn, Hg,
Mo, Ni, Au, Pt, Pb, W, Zn, Ca, Mg, Ba, Al, Si. Of these, Au, Ag, Cu, Fe, Pb, Sn,
Hg, S, and C were already known to the ancient world. The last decade of the eigh-
teenth century saw the discovery of U, Zr, Sr, Ti, Y, Be, Cr, and Te. The discovery
or isolation of the elements during the nineteenth century can be summarized as
follows:

1800–1809: Na, K, Nb, Rh, Pd,	1860–1869: Rb, In, Cs, Tl
Ce, Ta, Os, Ir	1870–1879: Sc, Ga, Sm, Ho,
1810–1819: Li, Se, Cd, I	Tm, Yb
1820–1829: Br, Th	1880–1889: Ge, Pr, Nd, Gd, Dy
1830–1839: V, La	1890–1899: He, Ne, Ar, Kr, Xe,
1840–1849: Ru, Tb, Er	Po, Ra, Ac
1850–1859:	

The data show an interesting pattern; during the first decade a considerable number of elements were discovered or isolated for the first time, reflecting the theoretical developments from Lavoisier and Dalton. The rate of discovery then fell off, until in the decade 1850–1859 no new elements were discovered. In 1860 Bunsen and Kirchoff demonstrated the utility of the spectroscope in the detection and identification of the elements, and in rapid succession Cs, Rb, Tl, and In were discovered with this new instrument. The 1870s and 1880s were marked by the discovery of several new lanthanides and the elements predicted by Mendeleev on the basis of the periodic table: eka-aluminum (Ga), eka-boron (Sc), and eka-silicon (Ge). In 1894 argon, the first of the inert gases, was discovered, followed rapidly by the remaining members of the group—Ne, He, Kr, Xe. The discovery of radioactivity by Becquerel in 1896 resulted in the recognition of polonium and radium by the Curies in 1898, and actinium by Debierne in 1899.

By 1900, therefore, the periodic table was essentially complete, except for some short-lived radioactive elements and for Eu (1901), Lu (1907), Hf (1923), and Re (1925). Proof of this, however, was first provided in 1914, when Moseley demonstrated the correlation between X-ray spectra and the atomic numbers of the elements.

Throughout the nineteenth century geochemical data were mainly the byproduct of general geological and mineralogical investigations and comprised more and better analyses of the various units—minerals, rocks, natural waters, and gases—making up the accessible parts of the earth. For many years this work was largely confined to European laboratories, but, with the organization of the U.S. Geological Survey and the appointment of F. W. Clarke as Chief Chemist in 1884, a center devoted to the chemical investigation of the earth was founded on the American continent.

Clarke was Chief Chemist for 41 years, until he retired in 1925. He was responsible for a vast and ever-growing output of analyses of minerals, rocks, and ores collected by the field staff or submitted for examination; moreover, he was always concerned with the fundamental significance of the mass of factual data thus acquired. In 1889 he published a classic paper, "The relative abundance of the chemical elements," which represented the first attempt to utilize the accumulated rock analyses to determine the average composition of the earth's crust and the relative abundances of the elements. It is interesting to note that even with the inadequate data then available, Clarke was able to draw some remarkably prescient conclusions. He wrote "If . . . .we assume that the elements have been evolved from one primordial form of matter, their relative abundance becomes suggestive . . . the process of evolution seems to have gone on slowly till oxygen was reached. At this point the process

exhibited its maximum energy, and beyond it the elements forming stable oxides were the most readily developed, and in the largest amounts. On this supposition the scarcity of the elements above iron becomes somewhat intelligible; but the theory does not account for everything and is to be regarded as merely tentative." However, in his great compendium, *The Data of Geochemistry,* which was first published as U.S. Geological Survey Bulletin 330 in 1908 and passed through five editions in less than twenty years, Clarke did not pursue this theory of the origin of the elements.

In many respects the appearance of the fifth edition of *The Data of Geochemistry* in 1924 marks the end of an era. During the preceding hundred years geochemical research was largely synonymous with the analysis of those parts of the earth accessible to visual inspection and chemical assay. From the nature of things it could be little more; interpretative geochemistry, the creation of a philosophy out of the mass of factual information, had to wait upon the development of the fundamental sciences, especially physics and chemistry. A single illustration serves to demonstrate this: the failure of all attempts to explain adequately the geochemistry of the silicate minerals before the discovery of X-ray diffraction provided a means for the determination of the atomic structure of solids.

The development of geochemistry in new directions was greatly advanced by the establishment of the Geophysical Laboratory by the Carnegie Institution of Washington in 1904. The policy followed in this laboratory of careful experimentation under controlled conditions, and the application of the principles of physical chemistry to geological processes, was an immense step forward. Previously geologists and chemists had been skeptical of the possibility of applying the techniques and principles of physics and chemistry to materials and processes as complex as those on and within the earth.

At the same time the original staff of the Geophysical Laboratory was beginning work in Washington, a new school of geochemistry was growing up in Norway. Fathered by J. H. L. Vogt and W. C. Brøgger, it attained worldwide distinction through the work of V. M. Goldschmidt and his associates. Goldschmidt graduated from the University of Oslo in 1911, and his doctor's thesis, *Die Kontaktmetamorphose im Kristianiagebiet,* was a basic contribution to geochemistry. It applied the phase rule, recently codified by the work of Bakhuis Roozeboom, to the mineralogical changes induced by contact metamorphism in shales, marls, and limestones, and it showed that these changes could be interpreted in terms of the principles of chemical equilibrium. During the next 10 years his work was devoted largely to similar studies on rock metamorphism. These studies stimulated related research in other Scandinavian countries and led eventually to the enunciation of the principle of "mineral facies" by Eskola in a paper published from Goldschmidt's laboratory.

In many ways 1912 can be considered a critical date in the development of geochemistry. In that year von Laue showed that the regular arrangement of atoms in crystals acts as a diffraction grating toward X-rays and thus made the discovery that enabled the atomic structure of solid substances to be determined. Since the geochemist is largely concerned with the chemistry of solids, the significance of this discovery can hardly be overestimated. However, some years elapsed before the impact of this new development was felt in geochemistry. In the fifth edition of *The Data of Geochemistry,* published in 1924, no mention is made of it. It is a tribute to Gold-

schmidt's insight that he not only realized the significance of crystal structure determinations for geochemistry, but also devised a plan of research that led to a maximum of results in a minimum of time. Between 1922 and 1926 he and his associates in the University of Oslo worked out the structures of many compounds and thereby established the extensive basis on which to found general laws governing the distribution of elements in crystalline substances. The results were published in a series of papers entitled *Geochemische Verteilungsgesetze der Elemente,* which in spite of the title dealt largely with the crystal structures of inorganic compounds. In these publications Goldschmidt's name is associated with T. Barth, W. H. Zachariasen, L. Thomassen, G. Lunde, I. Oftedahl, and others, all of whom have since had notable careers.

In 1929 Goldschmidt left Oslo for Göttingen, where he began investigations on the geochemistry of the individual elements, applying the principles discovered in the previous years and making use of the current development in quantitative spectrographic methods for rapidly determining small amounts of many elements with a high degree of precision. The results are well summarized in the seventh Hugo Müller lecture of the Chemical Society of London "The principles of distribution of chemical elements in minerals and rocks" (1937). Because of conditions in Germany, Goldschmidt returned to Oslo in 1935. Unfortunately, after the German invasion of Norway in 1940 he was able to do little more work. He was forced to flee Norway in 1942 to avoid deportation to Poland, and escaped to Sweden and thence to England. His health was seriously affected as a result of imprisonment in concentration camps in Norway, and he never recovered, dying in 1947 at the comparatively early age of 59.

Thanks largely to the work and the stimulus of Goldschmidt, the last 50 years have seen geochemistry develop from a somewhat incoherent collection of factual data to a philosophical science based on the concept of the geochemical cycle in which the individual elements play their part according to established principles. Geochemical speculation has extended beyond the accessible parts of the earth to the nature and constitution of the interior, the development of the earth throughout geological time, and ultimately to its pregeological history and to the history of the solar system as a whole.

An important school of geochemistry has developed in the U.S.S.R., especially since 1917. Its greatest names have been V. I. Vernadsky and his younger colleagues A. E. Fersman and A. P. Vinogradov. Its productivity has been immense. Geochemistry in the U.S.S.R. has been particularly directed toward the search for and exploitation of mineral raw materials, evidently with considerable success.

Most advances in geochemistry can be attributed to improvement in analytical techniques. The study of radioactivity during and after World War II led to enormous improvements in elemental analysis by neutron activation and opened up the opportunity for isotope studies and geochronometry by mass spectrometric techniques. Both of these powerful analytical methods continue to be improved and provide us with better data on smaller geochemical samples. This is important because it allows geochemical systems to be studied in even finer detail, leading to better control on geochemical hypotheses and theories.

The availability of lunar samples, remote probe data from planetary atmospheres

and surfaces and the acceptance of the concepts of plate tectonics in the 1960s and 1970s gave new perspectives to geochemical sampling and theory. The addition of extensive data from the moon as well as from the earth and meteorites allowed many theories of planetology and cosmochemistry to be tested. The concept of material recycling through subduction of ocean floor material required reevaluation of the geochemical cycles of the elements. Today, geochemistry is undoubtedly at the threshold of new and exciting discoveries and changes.

THE LITERATURE OF GEOCHEMISTRY

Geochemical facts, theories, and fantasies are widely scattered throughout the scientific literature. The first source of such data is, naturally, geological publications, but their number is legion and geochemical papers may be found in any of them. Chemical publications contain, of course, much that has geochemical significance, and occasionally classical material appears therein, as, for example, the symposium entitled "The physical chemistry of igneous rock formation," published by the Faraday Society in 1925. Fortunately, the chemical literature is relatively accessible through abstract journals. The literature of astronomy, biology, and of physics also carries information of importance to the geochemist.

Two journals devoted to geochemistry are *Geochimica et Cosmochimica Acta*, which began publication in 1950 under direction of an international board of editors, and *Geokhimiya*, published since 1956 by the Academy of Sciences of the U.S.S.R. The second has been published in translation by the Geochemical Society under the title *Geochemistry* (until 1963; in 1964 *Geochemistry* was replaced by *Geochemistry International*, comprising selected articles from *Geokhimiya* along with articles translated from other languages). Although these journals attract many specifically geochemical papers, many data of significance for geochemistry continue to appear as incidental material in publications concerned with mineralogy, petrology, economic geology, inorganic and physical chemistry, and other sciences. Three outstanding books are classic works in the field of geochemistry, one of which—the 1924 edition of Clarke's *The Data of Geochemistry*—is still a fine and readily available collection of analytical data on geological material. It is gradually being replaced by a new edition being published as separate chapters. The second is *Geochemistry* by Rankama and Sahama, published in 1950. This authoritative and comprehensive book not only treats the general aspects of the subject, but also gives a detailed account of the geochemistry of each element. The third is *Geochemistry* by V. M. Goldschmidt. Goldschmidt had prepared a great deal of manuscript for this work when he died in 1947, and thanks to the conscientious editorial work of Dr. A. Muir, and contributions to missing sections by other people, it was finally published in 1954. Although similar in scope to the book of Rankama and Sahama, it bears the distinctive stamp of Goldschmidt's genius for arriving at broad generalizations which correlate a large number of unconnected observations.

A more recent compilation of data, published in parts starting in 1969, is the *Handbook of Geochemistry* edited by K. H. Wedepohl. This useful handbook is produced so that individual parts may be corrected or replaced as new information becomes available.

SELECTED REFERENCES

Allègre, Claude-Jean, and G. Michard (1974). *Introduction to geochemistry.* 142 pp. D. Reidel Pub. Co., Dordrecht, Holland. A succinct review of geochemistry.

Brownlow, A. H. (1979). *Geochemistry.* 498 pp. Prentice-Hall, Englewood Cliffs, N.J. An introductory text in geochemistry.

Clarke, F. W. (1924). The data of geochemistry (fifth ed.). 841 pp. *U.S. Geol Surv. Bull.* 770.

Fairbridge, R. W. (ed.) (1972). *The encyclopedia of geochemistry and environmental sciences.* 1321 pp. Van Nostrand Reinhold Co., New York. Separate articles by selected authors on many aspects of geochemistry.

Fleischer, M. (ed.) (1962–). The data of geochemistry (sixth ed.). *U.S. Geol. Surv. Prof. Paper 440.* The new edition of this great work is a cooperative effort of many authors and is being published in parts.

Goldschmidt, V. M. (1954). *Geochemistry.* 730 pp. Clarendon Press, Oxford.

Krauskopf, K. B. (1979). *Introduction to geochemistry* (second ed.). 617 pp. McGraw-Hill Book Co., New York. The application of chemical principles to geological problems, useful tables.

Rankama, K., and T. G. Sahama (1950). *Geochemistry.* 912 pp. University of Chicago Press, Chicago.

Rösler, H. J., and H. Lange (1972). *Geochemical tables.* 468 pp. Elsevier Publishing Co., New York. A compendium of tables and diagrams, with explanatory text, covering all aspects of geochemistry.

Wedepohl, K. H. (ed.) (1969–78). *Handbook of Geochemistry* (5 vols.). Springer-Verlag, Berlin. Vol. 1 contains fundamental facts of geochemistry, geophysics, and cosmochemistry; the remaining volumes provide separate chapters dealing in a systematic format with the geochemistry of individual elements or groups of closely related elements. An important successor to Rankama's and Sahama's *Geochemistry.*

Weeks, M. E., and H. M. Leicester (1968). *Discovery of the elements* (seventh ed.). 910 pp. Chemical Education Publishing Co., Easton, Pa. A fund of interesting and significant data related to the history of geochemistry.

2 THE EARTH IN RELATION TO THE UNIVERSE

THE NATURE OF THE UNIVERSE

The earth is a unit within the solar system, which consists of the sun, the planets and their satellites, the asteroids, the comets, and the meteorites. The sun itself is only one star within our galaxy, which comprises probably more than 10^{11} stars and has a lens-like form with a diameter of about 70,000 light-years (1 light-year $\simeq 10^{13}$ km). Beyond our own galaxy there is a very large number of other systems of stars of approximately the same size. These systems, the extragalactic nebulae, are scattered fairly uniformly through space, the nearest to us being the Andromeda nebula at a distance of about 1.75×10^6 light years. The spectra of these extragalactic nebulae show a displacement of the lines toward the red end of the spectrum, this red shift being approximately proportional to their distance. The red shift is regarded as a Doppler effect due to recession of the nebulae with velocities approximately proportional to their distance, and leads to the picture of an expanding universe. A large part of the theory of the expanding universe is still highly speculative; for this reason conclusions based on the theory also partake of its speculative character.

THE AGE OF THE UNIVERSE

The theory of an expanding, dynamic universe implies that the universe has been and is in a state of evolution. If we extrapolate backward in time we arrive at a state when the universe was "contracted to a point," or when all the matter in the universe was concentrated into a very small region. It is customary to regard this as its primitive state and to reckon the astronomical age of the universe from this time. With certain assumptions as to the rate of expansion, the astronomical age of the universe is computed to be about 16×10^9 years.

The solar system can be treated as an isolated unit, and its age considered independently of the rest of the galaxy and the universe as a whole. The solar system is essentially a closed system, and its elemental composition the same as when it formed, except insofar as it has been modified by the conversion of hydrogen to helium and other nuclear reactions in the sun and by the decay of radioactive elements. The underlying assumption is that the material of the solar system was segregated at some definite time (the zero point of this time scale) and that the primitive constitution was subsequently modified by radioactive decay. The age of the elements is then reckoned from the time when the naturally radioactive series consisted entirely of the parent elements (nonradiogenic "daughter" elements may have been present). Some of the natural radioactive series of importance for this dating are given in Table 2.1.

Empirical and theoretical studies, as discussed later in this chapter, have shown that there are predictable regularities in the relative abundances of the elements and nuclides in the solar system. Thus predictions may be made of the original concentrations of some radioactive nuclides with respect to their neighbors. Since the abundance of ^{238}U with a half-life of 4.5×10^9 years and ^{235}U with a half-life of 7.1×10^8 years may be predicted to be about the same as that of the stable elements bismuth and mercury, the elements cannot have been formed more than several tens of thousand million years ago or the uranium isotopes would have decayed to levels difficult to detect. On the other hand the absence of ^{237}Np and ^{247}Cm, whose half-lives are 2.25×10^6 and 4×10^7 years, indicates that the formation of either of these nuclei occurred at least a few hundred million years ago, giving sufficient time for them to decay below detection levels. An independent solution is provided by consideration of the relative abundances of ^{235}U and ^{238}U. If these two isotopes were originally formed in approximately equal amounts, the present ratio of ^{235}U to ^{238}U (1:138) is due to the shorter half-life of ^{235}U, which causes it to decay much more rapidly than ^{238}U. The time required to reduce equal concentrations to the present ratio is about 6×10^9 years. Evidence of this sort, although qualified by uncertainties in the primordial abundances of the elements, indicates a limit to the age of the solar system.

Other arguments also support the view that the solar system dates from a few thousand million years ago. Thus of the isotopes of lead (^{204}Pb, ^{206}Pb, ^{207}Pb, ^{208}Pb), ^{204}Pb is not radiogenic, and its amount should be the same now as when it was orig-

Table 2.1 Some Natural Radioactive Series of Significance for Geological Dating

Parent Elements	Total Half-lives	End Products
^{238}U	4.5×10^9 years	$^{206}Pb + 8^4He$
^{235}U	7.1×10^8 years	$^{207}Pb + 7^4He$
^{232}Th	1.4×10^{10} years	$^{208}Pb + 6^4He$
^{87}Rb	4.9×10^{10} years	^{87}Sr
^{40}K	1.3×10^9 years	$^{40}Ar, {}^{40}Ca$

inally formed; the other isotopes are partly original and partly derived from the decay of ^{238}U, ^{235}U, and ^{232}Th. In material containing both uranium and lead the isotopic composition of the lead has undergone a progressive change during geological time; the relative amounts of the radiogenic isotopes have increased with respect to the nonradiogenic ^{204}Pb. Meteorites provide us with a particularly satisfactory illustration of this fact. Iron meteorites contain no uranium, and the small amount of lead present has the highest relative amount of ^{204}Pb of any natural material. Stony meteorites contain measurable amounts of uranium, and the lead present shows the effect of the continual addition of radiogenic lead in much higher $^{206}Pb/^{204}Pb$ and $^{207}Pb/^{204}Pb$ ratios than those for lead extracted from iron meteorites. A mathematical analysis of the data gives an age for meteorites of 4.6×10^9 years, which has been independently confirmed by rubidium-strontium age determinations on stony meteorites. This is the time since the iron and the stony meteorites were differentiated, and it can be plausibly equated with the time of planet formation in the solar system. Additional substantiation has come from the chronology determined on lunar samples, which shows that the moon aggregated 4.6×10^9 years ago.

Although the earth formed as an independent body in the solar system about 4.6×10^9 years ago, a long period evidently elapsed before the consolidation of the crust, judging from the information available from the dating of individual rocks of the crust by one or another of the radioactive decay schemes. On all the continents rock ages of around 2700 million years are well established. It has been much more difficult to extend the record further back, but the oldest ages from Africa are 3000 to 3600 million years; from North America 3100 to 3700 million years; from Europe 3500 million years and from Australia 3000 million years. The dynamic motions of the crust and evidence of a catastrophic meteoroid bombardment of the moon and earth about 4000 million years ago make it unlikely that the age record for crustal rocks will be extended much beyond 3700 million years. The time of formation of a primordial crust and the age of the earth as an individual body must, of course, be greater than these dates. The hiatus of several hundred million years between the age of the earth and that of the oldest crustal rocks was evidently occupied by the evolution of a relatively stable crust. Likewise dating of lunar rocks has found only a few with ages greater than 4000 million years.

THE NATURE OF THE SOLAR SYSTEM

In the study of geochemistry the solar system is of primary importance, although it is inconspicuous within our own galaxy and insignificant in relation to the universe as a whole. Data on the solar system are given in Table 2.2. Any satisfactory theory of the origin of the solar system must explain its regularities, the most important of which are the following:

1. The sun contains over 99.8% of the mass of the system, but only 2% of the angular momentum.
2. The planets all revolve in the same direction around the sun in elliptical orbits, and these orbits all lie in practically the same plane.

Table 2.2 Data on the Solar System

	Mass (earth = 1)	Radius (in kilometers)	Density (g/cm^3)	Surface Temperature (in K)	Surface Pressure (in bars)	Major and Minor Gases in Atmosphere
Sun	332,000	695,000	1.41	5500	—	H_2, He
Mercury	0.055	2440	5.44	620	—	—
Venus	0.815	6050	5.27	741	93	CO_2, N_2, H_2O, Ar, SO_2
Earth	1.00	6371	5.52	~290	1	N_2, O_2, CO_2, H_2O
Mars	0.11	3397	3.95	210–240	0.007	CO_2, N_2, Ar, O_2
Jupiter	318	71,600	1.31	170	High	H_2, He, CH_4, NH_3
Saturn	95.2	60,000	0.70	140	High	H_2, He, CH_4, NH_3
Uranus	14.6	25,900	1.21	80	High	H_2, He, CH_4, NH_3
Neptune	17.2	24,750	1.66	80	High	H_2, He, CH_4, NH_3
Pluto	0.0017	1300	1 ?	80	High	—

3. The planets themselves rotate about their axes in the same direction as their direction of revolution around the sun (except Uranus, and Venus, which have retrograde rotation); most of their satellites also revolve in the same direction.

4. The planets show a regular spacing as expressed by Bode's law,[1] and they form two contrasted groups: an inner group of small planets (Mercury, Venus, Earth, and Mars), which are called the terrestrial planets, and an outer group of large planets (Jupiter, Saturn, Uranus, and Neptune), which are called the major planets.

5. The major part of the angular momentum of the solar system is concentrated in the planets, not in the sun, in spite of the concentration of mass in the sun.

THE ORIGIN OF THE SOLAR SYSTEM

To be acceptable, any theory of the origin of the solar system must account for these regularities. Two main schools of thought exist, both with ancient and respectable antecedents. Both consider the solar system as derived from an ancestral sun or solar nebula. They differ essentially in that one prescribes the action of an external force

[1] Bode's law is an empirical series which closely approximates the relative distances of the planets from the sun and can be formulated as follows:

Mercury	Venus	Earth	Mars	(Vacant)	Jupiter	Saturn	Uranus	Neptune	Pluto
4	4	4	4	4	4	4	4		4
0	3	6	12	24	48	96	192		384
4	7	10	16	28	52	100	196		388

Actual distances of the planets from the sun in terms of the Earth's distance as 10

3.9	7.2	10	15.2		52	95	192	301	395

The law thus gives a good agreement with the measured distances, except for Neptune; the gap between Mars and Jupiter is occupied by the asteroids, at a mean distance of 29.

to form the planets from the sun, whereas the other rejects the idea of an external force and finds the energy required to form the planets within an ancestral solar nebula. The first school of thought dates back to the French philosopher Buffon in 1749, who suggested that the planets were torn from the body of the sun by a collision with another star. The other originated with the speculations of Kant in 1755. Kant suggested that within an original solar nebula, regions with slightly higher density than the mean would act as sinks for matter and the planets would thereby grow at the centers of these regions. Laplace in 1796 visualized the original state of the sun as a rotating tenuous mass of gas occupying the entire volume of the present solar system, from which contraction, accompanied by increasing rotational speed, led to the disengaging of a series of gaseous rings by centrifugal force, these rings then condensing to form the planets.

The Laplace hypothesis held the field for some 60 years, until the physicist Clerk Maxwell showed that the physics of the solar system is inconsistent with the mode of origin postulated by Laplace. The main difficulties lay in the concentration of angular momentum in the planets, and not in the sun, and in the mechanism by which the annular rings of gas might have condensed into planets. These considerations brought the Laplace theory into disfavor, and Buffon's theory was revived and given more precise form, first by Chamberlin and Moulton in the United States and later by Jeans and Jeffreys in England. These theories had in common the formation of the earth and the other planets from material torn from the sun by the impact or close approach of another star. The Chamberlin-Moulton hypothesis visualized the formation of the planets by the aggregation of small solid particles (planetesimals); the Jeans-Jeffreys hypothesis considered that the planets were formed by the condensation of masses of incandescent gas. These impact theories were favored for a number of years but have been abandoned since it was shown that material torn from the sun by the impact of another body could under no circumstances condense to form planets but would be completely dissipated throughout space in a very short time.

In recent years the origin of the earth and of the solar system as a whole has been the subject of intense speculation. Variants of the nebular theory have been suggested by a number of astronomers. The version originally proposed by von Weizsäcker seems to fit the facts best. It pictures the primitive sun as a rapidly rotating mass surrounded by an extended lens-shaped envelope consisting of solid particles and gas in turbulent motion. In this lens-shaped envelope eddy-like vortices would form, causing local accumulations of matter which aggregated to form the planets. Von Weizsäcker gives reasons for the regular spacing of these vortices, which would explain Bode's law. Von Weizsäcker's theory also explains the remarkable difference in size and density between the inner and the outer planets. In the lens-shaped envelope from which the planets condensed the temperature would decrease with distance from the center according to the inverse square law. Because of this falling-off of temperature more material could condense in the outer parts than in the inner parts. In the region of the inner planets only compounds of low volatility condensed, whereas in the outer regions the condensation products contained much material of low critical temperature. Because of this difference in the amount of condensation,

the outer planets also grew faster and larger than the inner ones. Hence the inner planets are small and dense, whereas the outer ones are large and have a low specific gravity. Pluto is exceptional but it may be an escaped moon of Neptune. Fine details in volatile abundances in the inner planets indicate that composition is not a simple function of heliocentric distance. After reviewing some differences between the compositions of the earth and Mars, Anders and Owen concluded that the larger inner planets attracted more late–stage volatile–rich material during their accretion than did the smaller Mars. Von Weizsäcker's theory incorporates part of the Chamberlin-Moulton hypothesis, in that it considers that the planets were built up by the aggregation of solid particles, that is, planetesimals, rather than by condensation of incandescent gas. This theory has been extensively developed by Urey, terHaar, and Kuiper; for a recent review see Hartman, 1972.

Other ideas on the mode of formation of the solar system are linked with the nature of double stars. All the double stars that have been observed have a large amount of angular momentum. Most of the single stars, such as the sun, have very little angular momentum, as far as we know. It is therefore reasonable to suppose that a double star may reach a condition of greater stability by evolving into a single star with a system of planets revolving about it, with the planetary system carrying the major portion of the angular momentum. Hoyle suggested that the evolution of a double star into a single star with a planetary system may result from the disintegration of one component of the double star with an accompanying supernova outburst. Most of the material of the supernova would be dispersed into space, but sufficient matter was left within the sun's sphere of influence to condense into the planets.

THE COMPOSITION OF THE UNIVERSE

Our knowledge of the chemical composition of the universe is obtained by spectroscopic examination of solar and stellar radiation, by the analysis of meteorites, and by what we know of the composition of the earth and other planets. Spectroscopic observation indicates the elements responsible for the radiation, and by careful analysis of the intensities of the spectral lines rough estimates can be made of the relative amounts of the different elements present in the outer layers of the radiating body. The data are consistent with the belief that the universe consists throughout of the same elements, and despite local variations which may generally be readily explained the relative abundances of the different elements are everywhere much the same. Only rarely has an element not previously known to occur on the earth been discovered elsewhere. Helium, first detected in the sun's spectrum by Lockyer in 1868, was identified on the earth by Ramsay in 1895 as the gas evolved when uraninite was heated with a mineral acid (Hillebrand a few years previously noticed this evolution of an inert gas from uraninite but thought it to be nitrogen). Technetium, an element not detected on the earth because of its short half-life, has been seen in some stellar spectra due apparently to its relatively recent synthesis in these stars.

THE COMPOSITION OF THE SUN

Spectroscopic studies of the sun have been made over many years and many data have been accumulated. The major limitations of this method of study are (a) some elements either do not give detectable spectra, or their strong lines have wavelengths less than 2900 Å and are absorbed by the atmosphere of the earth and cannot be observed (this limitation is being overcome by spectrographic data obtained at high altitudes by rockets and artificial satellites); (b) the spectra are produced in the outer part of the sun and give the composition of the solar atmosphere. Whether this composition is really representative for the sun as a whole depends on the effectiveness of convection to stir the material into a homogeneous mixture. These limitations must be borne in mind when considering the following information. About 70 elements have been recognized in the sun's spectrum, and there is no reason to conclude that any element is really absent; the presence of the others is unobservable because of their small abundance or the limitations previously stated. The relative abundances of the more common elements in the solar atmosphere are given in Table 2.3. The most striking feature is the extreme abundance of hydrogen and helium, which also make up most of the larger planets—Jupiter, Saturn, Neptune, and Uranus.

Table 2.3   Abundances of Elements in the Solar Atmosphere

Element	Atomic Number	Abundance (atoms/ 10^6 atoms Si)
H	1	2.2×10^{10}
He	2	1.4×10^{9}
C	6	9.3×10^{6}
N	7	2.0×10^{6}
O	8	1.6×10^{7}
Na	11	4.3×10^{4}
Mg	12	8.9×10^{5}
Al	13	7.4×10^{4}
Si	14	1.0×10^{6}
S	16	3.2×10^{5}
K	19	3.2×10^{3}
Ca	20	5.0×10^{4}
Sc	21	2.5×10^{1}
Ti	22	2.5×10^{3}
V	23	2.3×10^{2}
Cr	24	1.1×10^{4}
Mn	25	5.9×10^{3}
Fe	26	7.1×10^{5}
Co	27	1.8×10^{3}
Ni	28	4.3×10^{4}
Cu	29	2.6×10^{2}
Zn	30	6.3×10^{2}

After Ross and Aller, *Science,* **191,** 1223, 1976.

THE COMPOSITION OF THE PLANETS

Visual inspection and spectroscopic examination of the surfaces of the planets can tell us little about their bulk composition, since they are inhomogeneous, and their interior is undoubtedly different from their surface. However, data on their densities and analogies with the earth do provide some guide. Of the inner planets, Mercury has no atmosphere, and its density is similar to that of the earth. Venus is our nearest neighbor and has a very dense atmosphere, consisting almost entirely of carbon dioxide and nitrogen, which conceal its surface. The size and mass of Venus suggest that its composition is probably much like that of earth. Mars, the next planet beyond earth, has an atmosphere that does not obscure the surface of the planet and is therefore rarefied; however, clouds and dust storms have been observed on the face of Mars, and polar frost caps form in winter and disappear in summer. These polar caps appear to be formed of thin layers of H_2O ice, possibly with some solid CO_2 also. Much of the surface of Mars has a reddish or orange coloration, which has been plausibly ascribed to iron oxide coatings, similar to those on desert sands on earth. The size and mass indicate a bulk composition probably similar to that of earth. However, the oblateness of Mars and the lack of a magnetic field suggest that it does not have a fluid core.

Asteroids have not been sampled directly but spectral studies indicate compositions similar to meteorites. The albedo of Ceres and many other asteroids is very low, suggesting material of carbonaceous chondrite compositions. The densities of Ceres, Pallas, and Vesta have been determined as 2.1, 2.8, and 3.1, respectively, but the precision of the measurements is not high (about $\pm$ 0.5); however, these densities are within the observed range of carbonaceous chondrites and lower than those of most other meteorites.

The major planets, Jupiter, Saturn, Neptune, and Uranus, have many features in common, in particular low densities and thick atmospheres that completely obscure their surfaces. The low densities and thick atmospheres are explained by an abundance of hydrogen and helium probably comparable with that in the sun. Much of the hydrogen is evidently present as methane and ammonia. It has been shown that the rings of Saturn probably consist of ice particles, and the albedos and densities of some of the satellites of these planets suggest that they consist largely of ice also. The data we have on the major planets suggest that they have interiors chemically similar to that of earth, but they are covered with a great thickness of ice and condensed gases and have atmospheres containing hydrogen, helium, nitrogen, methane, and ammonia.

THE COMPOSITION OF METEORITES

Spectrographic evidence tells us nothing about the composition of the interiors of the planets. We must fall back on analogies with our own planet and with the evidence provided by meteorites, which are parts of the solar system (possibly fragments of disrupted asteroids) that eventually land on the earth. There are presumably millions of meteorites of all sizes in the solar system, from the finest dust particles up to those

that are miles in diameter (if we include the asteroids, which appear to be similar to meteorites in many respects). Meteoritic matter is continually falling on earth, mostly in the form of dust undetectable except by special means; it is estimated that the rate of meteoritic infall is between 30,000 and 150,000 tons annually. Our knowledge of the composition of meteorites comes from the larger and more spectacular ones that are seen to fall or from objects that are recognized as meteorites by the special characters distinguishing them from terrestrial rocks.

Meteorites consist essentially of a nickel-iron alloy, of crystalline silicate, mainly olivine or pyroxene, the iron sulfide mineral troilite, or of a mixture of these. No meteorite resembling sedimentary or metamorphic rocks has been found. Many systems of classification have been devised for meteorites, but for our purpose they may be grouped as follows:

1. Siderites or irons (average 98% metal).
2. Siderolites or stony irons (average 50% metal, 50% silicate).
3. Aerolites or stones.

The siderites, or iron meteorites, consist essentially of one or two nickel-iron metallic phases (Ni is usually between 4 and 20%, rarely greater), generally with accessory troilite (FeS), schreibersite (Fe, Ni, Co)$_3$P, and graphite. Additional accessory minerals, such as daubreelite (FeCr$_2$S$_4$), cohenite (Fe$_3$C), and chromite (FeCr$_2$O$_4$) occur more rarely. These accessory minerals are present as small rounded or lamellar grains scattered through the metal. The metal generally shows a definite structure known as Widmanstatten figures, which is brought out by etching a polished surface with an alcoholic solution of HNO$_3$. This structure consists of lamellae of kamacite (a nickel-iron alloy with about 6% Ni) bordered by taenite (a nickel-iron alloy with about 30% Ni). The lamellae are parallel to the octahedral planes of an originally homogeneous crystal of nickel-iron, and meteorites showing Widmanstatten structure are therefore known as octahedrites. This structure is typical of subsolidus exsolution in an alloy that has cooled very slowly from a high temperature. Hexahedrites are irons consisting entirely of kamacite, and ataxites are irons with more than 14% Ni and consist largely of taenite.

The siderolites, or stony-iron meteorites, are made up of nickel-iron and silicates in approximately equal amounts. Two distinct groups, the pallasites and the mesosiderites, of different chemical and mineralogical composition, are recognized. The pallasites are made up of a continuous base of nickel-iron enclosing grains of olivine which often show good crystal forms. In the mesosiderites the metal phase is discontinuous and the silicates are mainly plagioclase feldspar and pyroxene, sometimes with accessory olivine.

On the basis of texture the aerolites or stones are divided into two groups, the chondrites and the achondrites. The chondrites are so named because of the presence of chondrules or chondri, which are small rounded bodies (averaging 1 mm in diameter) consisting primarily of olivine and/or pyroxene. Chondrules seem to be unique to these meteorites and have never been observed in terrestrial rocks, hence are probably significant in terms of the origin of such meteorites. The average composition

of chondrites is about 40% olivine, 30% pyroxene, 5–20% nickel-iron, 10% plagioclase, and 6% troilite. One group of chondrites, the carbonaceous chondrites, is unique among meteorites in consisting largely of hydrated iron-magnesium silicate (serpentine or chlorite-like minerals) and containing up to 10% of complex organic compounds. The origin of these compounds, whether the remains of extraterrestrial organisms or the products of nonbiological synthesis, has been argued for over a century. However, recent research, especially on the Murchison meteorite which fell in Australia in 1969, strongly favors a nonbiological origin for the organic compounds. The Murchison meteorite contains a complex mixture of organic compounds (aliphatic and aromatic hydrocarbons, carboxylic acids, amino acids, and others). The similarity of these compounds to those produced by thermal synthesis from simple organic molecules supports a nonbiological origin, as does the nature of the amino acids, many of which are not found in natural proteins and none show optical activity.

The achondrites are a diverse group of stony meteorites that do not contain chondrules and are usually much more coarsely crystalline than the chondrites. Many achondrites resemble terrestrial igneous rocks in composition and texture, hence have probably crystallized from a silicate melt.

Tektites, enigmatic objects often considered with meteorites, consist of a silica-rich glass (average about 75% SiO_2) resembling obsidian, yet distinct from terrestrial obsidians in composition and texture. They have an unusual chemical composition, which consists of the conjunction of high silica and comparatively high alumina, potash, and lime with low magnesia and soda; this composition resembles a few granites and rhyolites, and some silica-rich sedimentary rocks. Tektites are found, generally as small (up to 200–300 g) rounded masses, in areas that preclude a volcanic origin. Unlike meteorites, tektites have not been observed to fall. Some authorities regard them as the product of the impact of comets or gigantic meteorites on the earth. In view of their enigmatic origin and their aberrant composition, they are not further considered here. A comprehensive account is provided by O'Keefe (1976).

Many chemical analyses have been made of meteorites. The irons form a rather homogeneous group, differing from one another essentially in nickel content, and an average composition is easily derived. The stony-irons and the achondrites are very diverse groups, and are comparatively rare, so averages of their compositions have little significance. The chondrites are abundant and show a remarkable homogeneity in chemical composition. Table 2.4 gives a selection of the available compositional data. The close correspondence between the average composition of iron meteorites with the average composition of the metal from chondrites strongly suggests a common source. The iron meteorites probably represent metal segregated by the partial or complete melting of material of chondritic composition.

There is general agreement that meteorites provide us with the best sample from which to derive the absolute abundances of the nonvolatile elements; several tables of elemental abundances have been compiled from the analytical data on meteorites. A principal difficulty in such compilations has been the selection of the analytical data, since meteorites differ greatly in composition and in relative abundance. This

Table 2.4   The Composition of Meteorite Matter

	Metal* (from irons)	Metal† (from chondrites)	Silicate† (from chondrites)	Average Chondrite†
	(weight per cent)			
O			43.7	33.24
Fe	90.78	90.72	9.88	27.24
Si			22.5	17.10
Mg			18.8	14.29
S				1.93
Ni	8.59	8.80		1.64
Ca			1.67	1.27
Al			1.60	1.22
Na			0.84	0.64
Cr			0.38	0.29
Mn			0.33	0.25
P			0.14	0.11
Co	0.63	0.48		0.09
K			0.11	0.08
Ti			0.08	0.06

*Brown and Patterson, *J. Geol.* 56, 87, 1948.
†Mason, *Amer. Museum Novitates,* No. 2223, 1965.

is shown in Table 2.5, which divides meteorites into two groups, the *finds* (those collected but not seen to fall) and the *falls* (those collected after having been seen to fall). The figures in Table 2.5 show a remarkable reversal in proportions between the finds and the falls. The reason is not far to seek. The relative abundance of irons as finds is due to their being easily recognized as meteorites, whereas a stony meteorite, unless seen to fall, could easily be overlooked as such. A truer indication of the relative abundance of the different meteorite types is therefore given by the relative proportions of those seen to fall. Such a compilation completely reverses the situation and indicates that the chondrites are far more abundant than all other types. The composition and structure of the chondrites favor the hypothesis that they may well represent fragments from planetesimals like those that aggregated to form the planets; the other meteorite types can plausibly be developed by the partial or complete melting and differentiation of material of chondritic composition. On this account the chemical composition of the chondrites has been the primary source of information regarding the absolute or cosmic abundances of the elements. The validity of this concept is illustrated in Figure 2.1, which compares solar abundances with those in the Allende carbonaceous chondrite for 19 elements, covering a wide range in abundance and chemical properties. Most of these abundances are identical within a factor of two which, in view of the inherent uncertainty in the solar abundances derived from spectrographic analysis, must be considered highly significant, indicating that chondrites are good average samples of solar system material.

Table 2.5   Frequency of Meteorite Finds and Falls

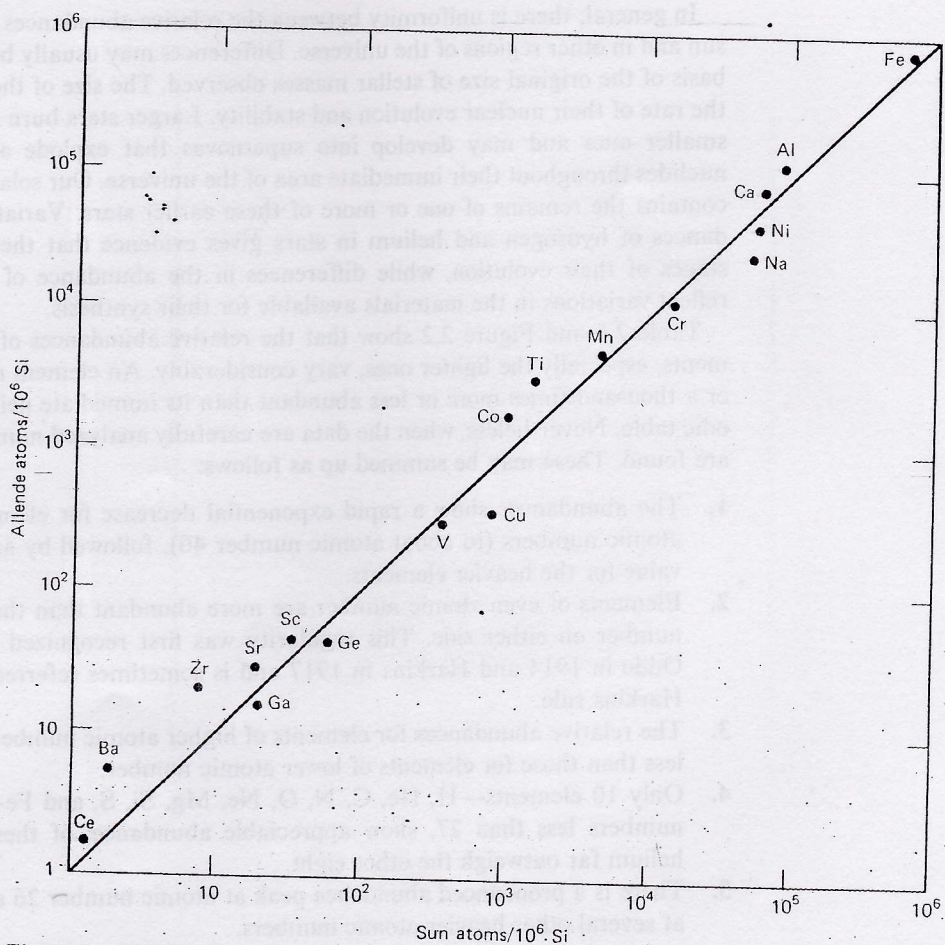| Type	Finds		Falls	
	Number	Percent	Number	Percent
Irons	545	58.1	33	4.6
Stony-irons	53	5.7	11	1.5
Achondrites	7	0.7	56	7.8
Chondrites	333	35.5	621	86.1
Total	938	100.0	721	100.0

Figure 2.1   Plot (logarithmic scale) of elemental abundances in the Allende meteorite versus solar abundances; the diagonal line represents equal abundances in the two materials (Mason, Accounts Chem. Res., 8, p. 218, 1975)

THE COSMIC ABUNDANCE OF THE ELEMENTS

On the basis of data on the composition of meteorites and of solar and stellar matter, Goldschmidt in 1937 compiled the first adequate table of cosmic abundances of elements and isotopes. The data on hydrogen and helium and other volatiles were derived largely from examination of the sun and stars, and the figures for most of the other elements were based on their relative abundances in meteoritic material. The overlap of nonvolatile abundant elements, such as silicon, is used to combine the data from both sources. Cameron has published a revised table (Table 2.6), using the more extensive and accurate data accumulated since 1937; the major features of Goldschmidt's abundance figures are not altered, although there are numerous differences in detail.

In general, there is uniformity between the relative abundances determined in the sun and in other regions of the universe. Differences may usually be explained on the basis of the original size of stellar masses observed. The size of the stars determines the rate of their nuclear evolution and stability. Larger stars burn more rapidly than smaller ones and may develop into supernovas that explode and disperse their nuclides throughout their immediate area of the universe. Our solar system probably contains the remains of one or more of these earlier stars. Variations in the abundances of hydrogen and helium in stars gives evidence that they are in different stages of their evolution, while differences in the abundance of heavier elements reflect variations in the materials available for their synthesis.

Table 2.6 and Figure 2.2 show that the relative abundances of the different elements, especially the lighter ones, vary considerably. An element may be a hundred or a thousand times more or less abundant than its immediate neighbor in the periodic table. Nevertheless, when the data are carefully analyzed numerous regularities are found. These may be summed up as follows:

1. The abundances show a rapid exponential decrease for elements of the lower atomic numbers (to about atomic number 40), followed by an almost constant value for the heavier elements.
2. Elements of even atomic number are more abundant than those of odd atomic number on either side. This regularity was first recognized independently by Oddo in 1914 and Harkins in 1917 and is sometimes referred to as the Oddo-Harkins rule.
3. The relative abundances for elements of higher atomic number than nickel vary less than those for elements of lower atomic number.
4. Only 10 elements—H, He, C, N, O, Ne, Mg, Si, S, and Fe—all with atomic numbers less than 27, show appreciable abundance; of these, hydrogen and helium far outweigh the other eight.
5. There is a pronounced abundance peak at atomic number 26 and smaller peaks at several other heavier atomic numbers.

The regularities displayed in Table 2.6 suggest that the absolute abundances of the elements depend on nuclear rather than chemical properties and are related to the inherent stability of the nuclei. An element is uniquely characterized by the num-

Table 2.6   Cosmic Abundances of the Elements in Atoms per 10^6 Atoms Si:*

Element	Abundance	Element	Abundance
1 H	2.66×10^{10}	44 Ru	1.9
2 He	1.8×10^9	45 Rb	0.40
3 Li	60	46 Pd	1.3
4 Be	1.2	47 Ag	0.46
5 B	9	48 Cd	1.55
6 C	1.11×10^7	49 In	0.19
7 N	2.31×10^6	50 Sn	3.7
8 O	1.84×10^7	51 Sb	0.31
9 F	780	52 Te	6.5
10 Ne	2.6×10^6	53 I	1.27
11 Na	6.0×10^4	54 Xe	5.84
12 Mg	1.06×10^6	55 Cs	0.39
13 Al	8.5×10^4	56 Ba	4.8
14 Si	1.00×10^6	57 La	0.37
15 P	6500	58 Ce	1.2
16 S	5.0×10^5	59 Pr	0.18
17 Cl	4740	60 Nd	0.79
18 Ar	1.06×10^5	62 Sm	0.24
19 K	3500	63 Eu	0.094
20 Ca	6.25×10^4	64 Gd	0.42
21 Sc	31	65 Tb	0.076
22 Ti	2400	66 Dy	0.37
23 V	254	67 Ho	0.092
24 Cr	1.27×10^4	68 Er	0.23
25 Mn	9300	69 Tm	0.035
26 Fe	9.0×10^5	70 Yb	0.20
27 Co	2200	71 Lu	0.035
28 Ni	4.78×10^4	72 Hf	0.17
29 Cu	540	73 Ta	0.020
30 Zn	1260	74 W	0.30
31 Ga	38	75 Re	0.051
32 Ge	117	76 Os	0.69
33 As	6.2	77 Ir	0.72
34 Se	67	78 Pt	1.41
35 Br	9.2	79 Au	0.21
36 Kr	41.3	80 Hg	0.21
37 Rb	6.1	81 Tl	0.19
38 Sr	22.9	82 Pb	2.6
39 Y	4.8	83 Bi	0.14
40 Zr	12	90 Th	0.045
41 Nb	0.9	92 U	0.027
42 Mo	4.0		

*A. G. W. Cameron, written communication, 1980.
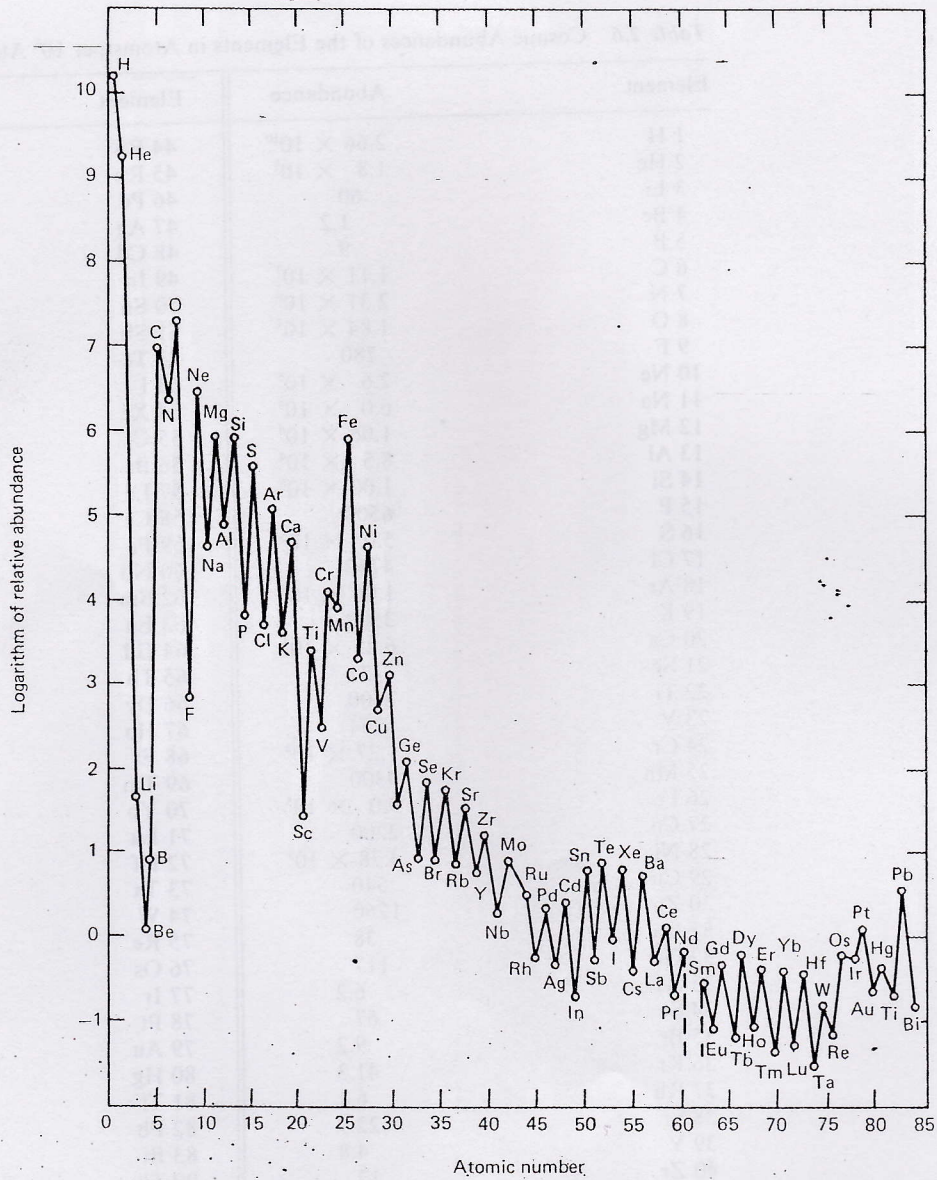
Figure 2.2 Relative abundances of the elements, referred to Si = 10^6, plotted against atomic number.

Table 2.7 Illustration of Isotopes, Isobars, and Isotones

Isotopes $Z = 20$ (calcium)				Isotones $N = 20$			
N	A	Percent of Element	Element		Z	A	Percent of Element
20	40	96.97	Sulphur		16	36	0.0136
22	42	0.64	Chlorine		17	37	24.471
23	43	0.145	Argon		18	38	0.063
22	44	2.06	Potassium		19	39	93.10
26	46	0.0033	Calcium		20	40	96.97
28	48	0.185					

Isobars $A = 40$			
Element	Z	N	Percent of Element
Argon	18	22	99.61
Potassium	19	21	0.0119
Calcium	20	20	96.97

ber of protons (Z) in its nucleus, but the number of neutrons (N) associated with these protons can vary. As a result, an element may have several *isotopes* differing in mass number or atomic weight $A(A = N + Z)$ and stability but not appreciably in chemical properties. Similarly, there are *isobars,* which are different elements with the same A but different values of N and Z, and *isotones,* which are different elements with the same neutron number N but with different values of A and Z (Table 2.7).

Relatively few of the possible nuclides of any element are stable; of the thousand or more isotopes known to date, only about 270 are nonradioactive. The implication is that a nuclide is abundant because the combination of protons and neutrons in its nucleus is particularly stable. On this basis the drop in relative abundances with increasing nuclear complexity can readily be explained; the absence from the earth of elements 43, 61, 85, and 87 is due to an almost complete instability of any nuclear arrangement for these atomic numbers (note that all four are odd-numbered). Then, too, it must be more than a coincidence that the nuclei of lithium, beryllium, and boron, which are exceptionally rare among low-numbered elements, are just those nuclei that are most readily disintegrated by bombardment with protons, alpha particles, and neutrons. As might be expected, the breakdown of relative abundance data for the elements into isotopic abundances has led to some significant results. Nuclei of the even N (neutron number)-even Z (proton number) type are both more numerous and more abundant than any of the other types. Nuclei of the even N-odd Z and odd N-even Z are about equally numerous and abundant. Nuclei of the odd N-odd Z type are few in number and low in relative abundance except for ^{14}N. These features evidently reflect the nuclear binding energy, which is greatest for even N-even Z nuclei.

THE ORIGIN OF THE ELEMENTS

The structure of the nuclei of the elements as aggregates of protons and neutrons has resulted in theories to explain their origin and their relative abundances by a synthesis, or buildup, starting with either or both of these basic building blocks.

Several theories as to the mode of formation of the chemical elements have been proposed. One, which may be termed the equilibrium theory, proposes that the relative abundances of the elements are the result of a "frozen" thermodynamic equilibrium between atomic nuclei at some high temperature and density. By suitable assumptions as to the temperature, pressure, and density, good agreement with observed abundances is obtained for elements of atomic number up to 40. For elements of higher atomic number, however, these assumptions lead to impossibly low abundances. On this account, theories have been proposed that consider the relative abundances of the elements as resulting from nonequilibrium processes; on this basis the light nuclei were built up by thermonuclear processes and the remaining nuclei by successive neutron capture, with intervening β^- disintegrations. This theory predicts the general trend of the observed data but fails to explain some of the detailed features, particularly bridging of the gap caused by the nonexistence of nuclei of atomic weights 5 and 8. It is apparent that no single process can satisfactorily account for the observed complexities.

Advances in nuclear physics allowed Burbidge, Burbidge, Fowler, and Hoyle in 1957 to propose that the general features of the abundance curve could be explained by nuclear reactions taking place in stars. In order to explain the measured distributions they outlined eight processes for the synthesis. In their model, elemental matter started with hydrogen, which formed the primitive matter from which stars were made. The individual steps in their synthesis are:

1. *Hydrogen "burning"* to produce helium. Hydrogen burns by successive proton capture to produce ^{3}He. At temperatures of 10^7°C and densities about 100 g/cm^3, ^{3}He nuclei impact and combine to form ^{4}He and two protons. At still higher temperatures additional nuclear reactions involving ^{3}He, ^{4}He, and protons provide alternate mechanisms for producing additional ^{4}He. In older or second generation stars containing ^{12}C a catalytic cycle progressing through six steps allows the overall reaction $4H^1 \rightarrow He^4$ to take place, as follows:

   $^1H + {}^+H \rightarrow {}^2D + \beta^+ + \nu$ $\qquad$ $^{12}C + {}^1H \rightarrow {}^{13}N + \gamma$ (gamma ray)

   $^2D + {}^1H \rightarrow {}^3He + \gamma$ $\qquad$ $^{13}N \rightarrow {}^{13}C + \beta^+$ (positron) $+ \nu$ (neutrino)

   $^3He + {}^3He \rightarrow {}^4He + 2{}^1H$ $\qquad$ $^{13}C + {}^1H \rightarrow {}^{14}N + \gamma$

   $\qquad\qquad\qquad\qquad\qquad\qquad$ $^{14}N + {}^1H \rightarrow {}^{15}O + \gamma$

   $\qquad\qquad\qquad\qquad\qquad\qquad$ $^{15}O \rightarrow {}^{15}N + \beta^+ + \nu$

   $\qquad\qquad\qquad\qquad\qquad\qquad$ $^{15}N + {}^1H \rightarrow {}^{12}C + {}^4He$
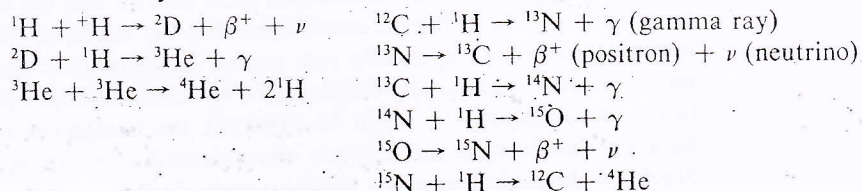
   The C-cycle reaction also provides a mechanism for the production of quantities of ^{14}N in the star.

2. *Helium burning* by three body collisions or "triple α" reactions takes place at temperatures in the order of 10^8°C and densities of about 10^5 g/cm^3. This process provides a mechanism for skipping the unstable nuclides with masses of 5

and 8. The ^{12}C produced can then add ^4He to produce ^{16}O. At this point in a star's evolution, ^1H has been converted to ^4He and the ^4He into ^{12}C and ^{16}O with the accumulation of some ^{14}N.

3. Further collapse of the star producing higher temperatures and greater densities allows additional reactions, sometimes known as the α *process* or carbon and oxygen burning, to take place. In this stage of steller evolution ^{12}C nuclei react to form species such as ^{20}Ne, ^{23}Na, ^{23}Mg, and ^{24}Mg together with a new supply of α particles and protons. In sequence ^{16}O nuclei react to produce ^{28}Si, ^{31}P, ^{31}S, and ^{32}S and probably some Cl and Ar from secondary reactions involving the α particles, protons, and neutrons.

4. As the star evolves further, silicon burning, or the equilibrium *e-process*, takes place. Because of the high temperatures, the rate of nuclear reactions is increased and elements with atomic numbers greater than ^{28}Si come into equilibrium with it. This process is responsible for the nucleosynthesis of the most abundant nuclei between $A = 28$ and $A = 57$ at the Fe peak. With the production of the Fe-group elements the star will have run out of reactions that can supply it with energy, since ^{56}Fe is at the top of the curve of binding energy per nucleon.

5. The primary mechanism for the synthesis of elements heavier than iron is by the capture of free neutrons. Neutrons are apparently produced by the reactions of α particles on ^{13}C, ^{17}O, or ^{21}Ne. Slow neutron capture, or the *s-process*, can produce elements up to and including ^{209}Bi, but many observed nuclei are bypassed. The process is called the *s*-process because the rate of neutron addition is slow compared with beta-decay lifetimes of the nuclei produced. Any unstable nuclei produced generally decay before the next neutron interaction takes place. The isotopes of a given element are built up successively until one is reached that is unstable. Then β decay produces the isobaric nuclide of that element which serves as the target for further isotope building reactions until another unstable species is formed. This process produces peaks in the element abundance curve where there is a build-up of stable elements with low neutron capture cross-sections.

6. To produce elements heavier than bismuth and to make neutron-rich elements not made by the *s*-process, a rapid neutron capture, or *r-process*, has been proposed. In this process neutrons are added to β–decay unstable nuclei before the decay can take place. This allows the production of many nuclei, which may through subsequent decay account for the neutron-rich isotopes of some elements. The large neutron fluxes for the *r*-process appear to take place during the massive stellar explosions called supernovas. These explosions provide the large flux of neutrons required and a mechanism for dispersing the elements made by all of the processes into space so that they may be recycled.

7. The relatively rare proton-rich isotopes of the heavy elements not produced by the neutron capture processes are explained by a rapid proton capture or *p-process*. Such a process may take place in the outer parts of a supernova explosion. Figure 2.3 illustrates the proposed mode of origin of nuclides in a selected portion of the chart of the nuclides.

Element–atomic number	46	47	48	49	50	51	52	53	54	55	56	57	58	59
42–Mo					15.8		94 / 9%	95 / 15.7%	96 / 16.5%	97 / 9.5%	98 / 23.8%	99 / 66h	100 / 9.6%	
41–Nb							93 / 100%	94 / 2×10^4 y	95 / 35d					
40–Zr					90 / 51.4%	91 / 11.2%	92 / 17.1%	93 / 10^4 y	94 / 17.4%	95 / 65d	96 / 2.8%			
39–Y					89 / 100%	90 / 64h								
38–Sr	84 / 0.6%		86 / 9.9%	87 / 7%	88 / 82.6%	89 / 51d								
37–Rb			85 / 72.2%	86 / 18.6d	87 / 27.8%									

Neutron number
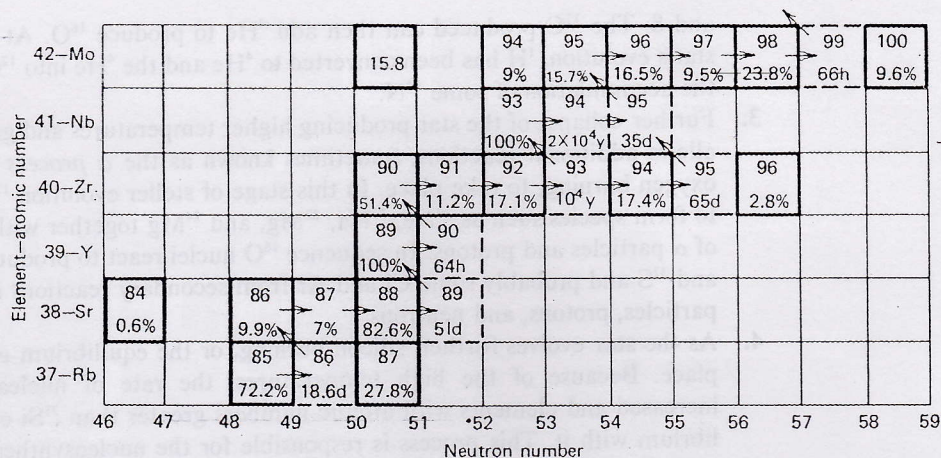
Figure 2.3   Portion of the chart of the nuclides. The fraction of stable isotopes is shown for each element, together with an *s*-process formation path indicated by arrows. Nuclides to the left of the *s*-process nuclides are formed by the *p*-process and those to the right by the *r*-process. Some nuclides may be formed by more than one process.

8.   The production of Li, Be, and B not explained by the above mechanisms has been attributed to an *x-process,* most likely as spallation products of lightweight abundant nuclides of carbon or oxygen.

These processes are correlated with the observed features of stellar evolution. All stars convert hydrogen into helium, but only the most massive stars produce the elements in the upper part of the periodic table. Certain heavy nuclides appear to be formed only under catastrophic conditions, such as the development of a supernova. A supernova is essentially a stellar explosion, the catastrophic disintegration of a star. The explosion produces luminosity of the order of 10^8 that of the sun, and the luminosity falls off exponentially with a half-life of about 56 days. It can hardly be a coincidence that ^{254}Cf decays by spontaneous fission with a half-life of 56 days. Evidently a supernova is triggered by the *r*-process. Man has reproduced the *r*-process on a comparatively modest scale; substantial quantities of californium are produced in H-bomb explosions when the ^{238}U in the bomb is exposed to an intense neutron flux during the explosion.

SELECTED REFERENCES

Aller, L. H. (1961). *The abundance of the elements.* 283 pp. Interscience Publishers, New York. A comprehensive account of elemental abundances, with special reference to solar and stellar abundances and an excellent discussion of theories of element synthesis.

Anders, E. (1977). Chemical compositions of the Moon, Earth, and eucrite parent body. *Phil. Trans R. Soc. Lond. A.* **285**, 23–40. A systematic approach to the estimation of planetary compositions.

Anders, E. and T. Owen (1977). Mars and Earth: Origin and abundance of volatiles. *Science* **198**, 453–465. An important paper giving evidence for the accretion of planetary bodies with differing amounts of different primary components.

Hartmann, W. K. (1972). *Moons and planets: an introduction to planetary science.* 404 pp. Bogden & Quigley, Inc., Belmont, Ca. A comprehensive account of the solar system.

King, E. (1976). *Space geology.* 349 pp. John Wiley & Sons, New York. A geological review of meteorites, the moon, and other planetary bodies.

Kuiper, G. P. (ed.) (1953–1963). *The solar system* (4 vols.). University of Chicago Press, Chicago. A comprehensive reference work by many contributors; Vol. 1 deals with the sun; Vol. 2, with the earth; Vol. 3, with planets and satellites; and Vol. 4, with the moon, meteorites, and comets.

Larimer, J. W. (1978). Meteorites: Relics from the early solar system. In *The origin of the solar system,* S. F. Dermott (ed.), John Wiley & Sons, New York, pp. 347–393.

O'Keefe, J. A. (1976). *Tektites and their origin.* 254 pp. Elsevier Scientific Publishing Co., New York.

Sears, D. W. (1978). *The nature and origin of meteorites.* 187 pp. Oxford University Press, New York. A summary of many aspects of meteorites.

Taylor, S. R. (1975). *Lunar science: A post-apollo view.* 372 pp. A compilation and review of the results of the Apollo missions.

Wood, J. A. (1979). *The solar system.* 196 pp. Prentice-Hall, Englewood Cliffs, N.J. A review of the solar system and its origin including meteorites.

3 THE STRUCTURE AND COMPOSITION OF THE EARTH

INTRODUCTION

A knowledge of the composition and state of the earth's interior is prerequisite to an understanding of geochemistry. This problem obviously cannot be solved by direct observation. The deepest boring yet made is one of about 10 km; mines are still shallower; and apart from borings and mines the only accessible parts of the earth are those exposed on the surface itself. Some material is brought to the surface by igneous activity, but as we know only incompletely the depth from which it comes and whether it has undergone any second-order changes, few deductions can be made from this evidence. Thus to obtain some understanding of the earth's internal structure we must use an indirect approach. Here we turn to geophysics, the application to the properties of the earth and its several parts of physical laws dealing with gravitation, wave transmission, heat conduction, and other phenomena. The principal sources of information are (a) the acceleration of gravity at the earth's surface and the gravitational constant, from which the mean density of the earth can be determined; (b) the constant of precession of the equinoxes, from which the earth's moment of inertia can be calculated, thereby allowing important inferences to be drawn regarding density distribution within the earth; (c) seismological data, which indicate the presence of discontinuities within the earth and from which information can be derived on the elastic constants of the materials in the interior; (d) heat flow data, which reflect the abundance and distribution of radioactive elements in the crust and mantle. These facts, together with laboratory determinations of the elastic constants of various rocks, information on the probable abundances of the elements, and similar data, provide the basis for theories regarding the earth's internal structure and composition. To be acceptable any theory must be consistent with the data available. However, the data may permit of several interpretations, none of which

28

will formally conflict with what is known about the earth. Current ideas on the earth's internal constitution are certain to be modified by the discovery of new facts and the improvement of existing knowledge.

SEISMIC DATA ON THE EARTH'S INTERIOR

Much information about the earth's interior is derived from the analysis of earthquake waves. An earthquake generates waves of various kinds, of which the two types that pass through the body of the earth are the most important for our present purpose. These two types of waves travel with unequal velocity, even in the same medium. The faster are those transmitted by vibrations in the direction of propagation (analogous to sound waves in air). They are the first to be recorded by seismographs at an appreciable distance from the focus of the earthquake and are called primary or P waves. The slower waves are transmitted by vibrations at right angles to the direction of propagation (analogous to light waves) and are known as secondary, shear, or S waves. The velocities of P and S waves vary with the density and elastic constants of the material through which they pass, and they are subject to reflection and refraction at surfaces of discontinuity. By comparing the times at which P and S waves from the same shock arrive at different stations, travel-time tables can be drawn up from which the velocity of these waves as a function of depth can be calculated. Figure 3.1 depicts graphically the data thus obtained. It shows that the interior of the earth is clearly heterogeneous, in the sense that at different depths the material has different elastic properties. This heterogeneity is not random but is present in zones separated from each other by discontinuities of greater or lesser sharpness. Two major or first-order discontinuities (a first-order discontinuity is one producing an abrupt break in the velocity-depth curve) have been recognized
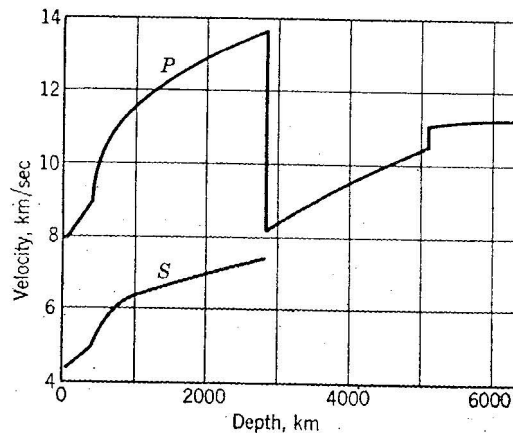
Figure 3.1 Velocity of P and S waves within the earth. (After Bullen, *An introduction to the theory of seismology.* Courtesy of Cambridge University Press)

by all geophysicists. The earth is thus divided into three parts: the crust, from the surface down to the first discontinuity (the Mohorovičić discontinuity); the mantle, from the base of the crust to the second discontinuity (the Wiechert-Gutenberg discontinuity); and the core, from the Wiechert-Gutenberg discontinuity to the center of the earth. The extinction of the S waves at the base of the mantle is particularly significant, suggesting that the material of the underlying core lacks rigidity and behaves as a liquid. Second-order discontinuities are recognized within the crust, the mantle, and the core (a second-order discontinuity is marked by a sudden change in the *rate* at which wave velocity increases or decreases and produces a change in slope of the velocity-depth curve). There is no general agreement on the number and positions of these discontinuities, and their interpretation is much less certain.

DENSITY WITHIN THE EARTH

It may be said that Sir Isaac Newton was the founder of geophysics because his formulation of the law of gravitation provided the means for determining the mass of the earth and its mean density. Newton made the prescient statement: "It is probable that the quantity of the whole matter of the earth may be five or six times greater than if it consisted of water." It was not until the latter part of the eighteenth century that Newton's brilliant guess was confirmed by experiment. In 1798 Cavendish determined the constant of gravitation by comparing the attraction between two lead spheres and two gold spheres, using a sensitive torsion balance; from this measurement he derived the figure of 5.48 for the mean density of the earth. The present accepted figure is 5.517 ± 0.004. Since the mean density of the surface rocks is about 2.8, it follows that the interior must have a density greater than 5.5 to account for that of the Earth as a whole. The high density may be explained in two ways:

1. A change in physical state, the increase in density being due to the contraction of crustal material into much smaller volume under enormous pressure.
2. A change in chemical composition, the increase in density then being due to the presence of some intrinsically heavier substance, such as a heavy metal.

The next step is the determination of the density distribution within the earth. It was mentioned in the previous section that the velocities of P and S waves vary with the density and elastic constants of the material through which they pass. The relevant equations are

$$V_p^2 = \left(\frac{1}{\rho}\right)\left(k + \frac{4\mu}{3}\right)$$
$$V_s^2 = \frac{\mu}{\rho}$$

(V_p and V_s are the velocities of P and S waves, respectively; ρ is the density of the material, k the bulk modulus, μ the rigidity.) Of the variables in these equations only the velocities are accurately known for conditions in the earth's interior, and thus no unique solution for the other factors is possible from seismic data alone. However,

the density distribution within the earth must comply with two stringent conditions: the integrated density must agree with the known density of the earth as a whole, and it must also give the correct moment of inertia as determined from the precession of the equinoxes. By making some plausible assumptions in the interpretation of the seismic data, Bullen computed figures for the density distribution within the earth which are consistent with the independent controls of bulk density and moment of inertia (Figure 3.2). From these results he also calculated the pressure distribution within the earth (Figure 3.3). The pressure at the earth's center is found to be 3640 kilobars, or somewhat over 3,000,000 atm. Later work has shown that the generalized curves constructed by Bullen show significant fine structure especially in the upper mantle. Such structure is shown in Figure 3.4 together with an interpretation of its causes.

Region a is the lithosphere, a slab of rock material of variable thickness. The crust forms the upper part of this unit, and it includes the Mohorovičić discontinuity. Its lower boundary is marked by a region of decreasing S wave velocity. The lithosphere is rigid and forms the numerous plates that drift over the earth's surface.

Region b is the asthenosphere, a seismic-shear-wave low velocity zone, which is interpreted as a zone of partial melting. Regions a and b do not appear to have greatly different compositions, but their physical states differ significantly. However, isotopic studies of basalts do show that those erupted at midocean ridges have come from regions chemically different from those erupted at other areas on the earth's surface, including some oceanic islands and continental flows. Based on isotopic studies of the systems Sm-Nd and Rb-Sr, it has been shown that the midocean ridge basalts are derived from an upper mantle source region depleted in those elements commonly found in the crust whereas continental basalts come from a deeper region undepleted in these materials. The upper mantle is a complex heterogeneous unit.
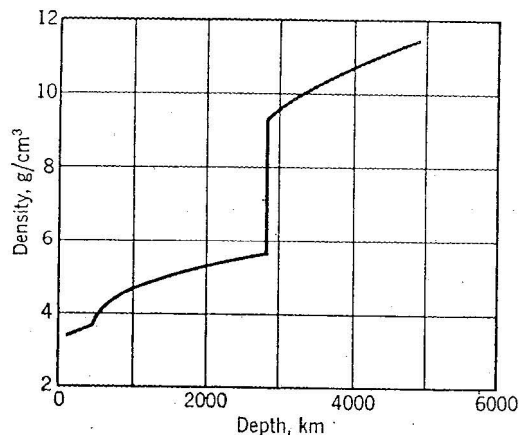
Figure 3.2   Density variation in the earth's interior. (From Bullen, *An introduction to the theory of seismology*. Courtesy of Cambridge University Press)
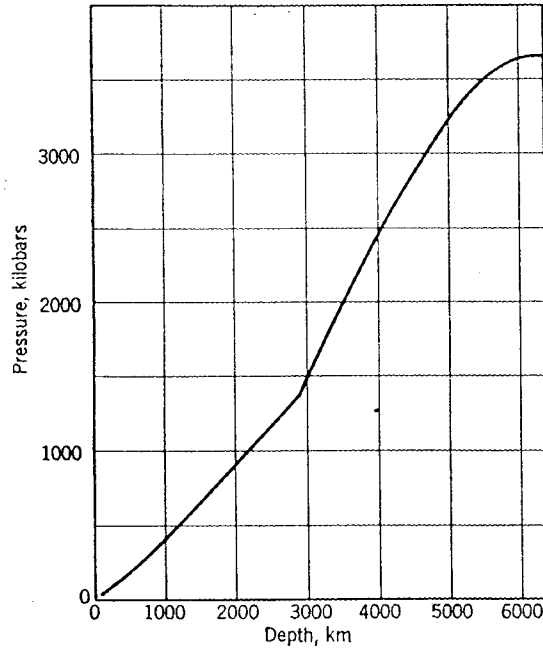
Figure 3.3 Pressure distribution in the earth's interior. (From Bullen, *An introduction to the theory of seismology*. Courtesy of Cambridge University Press)
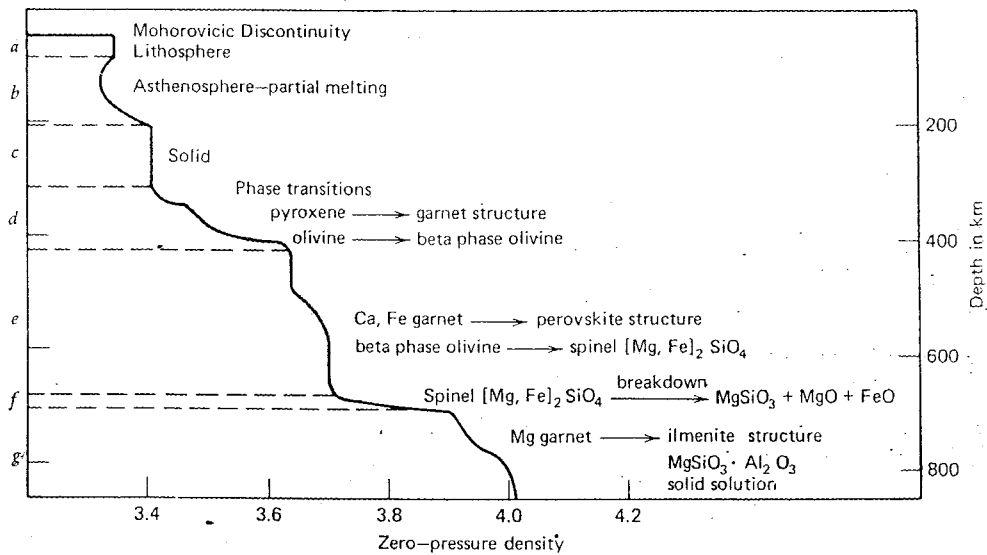
Figure 3.4 Phase assemblages with changing zero-pressure densities for a model of the upper mantle. (After Ringwood, Origin of the Earth and Moon. Courtesy of Springer-Verlag)

Below the asthenosphere the rocks again become solid and rigid with only slight velocity and density change with depth. In the region of 300- to 400-km depth a rapid increase in density and velocity takes place. Based on theory and experiment, Ringwood and his co-workers have shown that at these pressures phase transformations take place. Pyroxene assumes a garnet structure, and olivine transforms to the beta Mg_2SiO_4 structure, which is about 8 percent denser than olivine.

At greater depths in the e region (below 400 km), calcium garnet transforms to the dense perovskite structure and β-Mg_2SiO_4 changes to a spinel structure. Below this, in the f and g regions, the spinel and garnet phases begin to break down into simpler dense components. Spinel changes to MgO with a rock salt structure and to $MgSiO_3$ with a perovskite structure, and garnet changes to the ilmenite structure followed by another change to the perovskite structure. All of these phase changes on essentially a constant chemical composition can explain the observed physical properties.

TEMPERATURES WITHIN THE EARTH

We know the variation of density and pressure within the earth to a considerable degree of preeision, but estimates of temperatures at depth are little more than intelligent guesses. Direct observations in mines and boreholes show that temperature increases with depth, although the rate of increase varies greatly from place to place. This variation is characteristic of the crustal layers, and at quite moderate depths the temperature gradient probably becomes uniform. Measured thermal gradients in the crust range from 10 to 50°/km, and an average value of 30°/km is often used. From the thermal gradient and the thermal conductivity of the rocks the flow of heat toward the surface can be calculated. The average value of this heat flow is 1.5×10^{-6} cal/(cm²)(sec). This is about 50 cal/cm² annually, sufficient to melt a sheet of ice 6 mm thick (latent heat of fusion of ice is 80 cal/g). This shows how little the earth's internal heat can influence climatic conditions; 6 mm of ice can easily be melted away by a few hours' sunshine or be formed in one frosty night. However, when computed for the area of the earth it is nevertheless much larger than the amount of heat brought to the surface by the spectacular activity of volcanoes.

Selected heat flow values for different geologic regions are given in Table 3.1. It is significant that the average oceanic heat flow is approximately the same as the average continental heat flow. When first noted this was a surprising discovery, because a considerable fraction of the heat flow from the continents is accounted for by the radioactive elements in the continental crust, but the oceanic crust is thin, less radioactive, and incapable of providing more than 10% of the oceanic heat flow. Using data on the distribution of radioactive elements in different rock types it is possible to estimate the contribution of radioactive heat from crustal rocks (Table 3.2) and thus determine the heat concentration from the mantle. Such calculations indicate that crustal rocks may contribute from 30 to 50% of the continental heat flow. The remainder of the continental heat flow and the oceanic heat flow is deep heat from the earth's interior.

Table 3.1 Heat-Flow Values for Major Physiographic Provinces

Province	Number of Measurements	Mean Heat Flow q (μcal/(cm^2)(sec)	Standard Deviation
Ocean basins	273	1.28	0.53
Ocean ridges	338	1.82	1.56
Ocean trenches	21	0.99	0.61
Other seas	281	1.71	1.05
Precambrian shields	26	0.92	0.17
Phanerozoic nonorogenic	23	1.54	0.38
Phanerozoic orogenic areas	68	1.48	0.56
Paleozoic orogenic areas	21	1.23	0.40
Mesozoic-Cenozoic orogenic	19	1.92	0.49
Island arc areas	28	1.36	0.54
Cenozoic volcanic areas	11	2.16	0.46

After Lee and Uyeda, 1965; "Review of Heat Flow Data" in *Terrestrial Heat Flow*, W. H. K. Lee (ed.), Geophysical Monograph 8, American Geophysical Union.

Table 3.2 Heat Production by Igneous Rocks

Type of Rock	Heat Produced by U, (ergs/g year)	Heat Produced by Th, (ergs/g year)	Heat Produced by K, (ergs/g year)	Total Heat Production, (ergs/g year)
Granites	117	84	34	235
Acidic	126	109	38	273
Intermediate	43	36	29	108
Intermediate	81	81	29	191
Basalts	25	41	6.4	72
Dunites	0.42	0.44	0.01	0.87

From Jacobs, Russell, and Wilson *Physics and Geology*, Second Edition (New York: McGraw-Hill, © 1974).

Ocean ridges have on the average higher heat flows than do ocean basins which in turn have higher average heat flows than do ocean trenches. This is interpreted to indicate and support the idea that significant ocean floor heat is carried to the surface by large mantle convection cells involved in tectonics. Such cells surfacing at ocean ridges provide the higher heat flows observed and also may explain the ocean floor subsidence away from the ridges due to cooling and contraction of the ocean floor crust.

Seismic evidence shows that the mantle is solid throughout, except perhaps in some parts of the asthenosphere, and therefore the temperature at any depth within it cannot exceed the melting range of the material at that depth. Data on the variation of melting point with pressure for silicates are not complete, but for diopside (Figure 3.5) about 10°/kilobar is indicated, or 3°/km. On this basis the temperature at the bottom of the mantle cannot be greater than about 10,000°. However, the

melting point curves for the silicates in Figure 3.5 have a pronounced curvature, and their slopes decrease markedly with increasing pressure, so that the melting points at great depths will be considerably lower than given by the preceding extrapolation. The melting point of the material at the base of the mantle has been estimated at about 4000°. This is consistent with a molten iron core, because the melting temperature of iron at the pressure of the core-mantle boundary is believed to be in this range.

Figure 3.6 illustrates an estimated geotherm or thermal gradient within the earth and the melting relationships of mantle and core material. Note the close approach of the geotherm to the silicate melting point curve in the upper mantle where basaltic magmas are generated and in the outer core which appears to be liquid nickel-iron. Such curves are somewhat speculative but probably represent the thermal conditions in the earth to the first approximation.
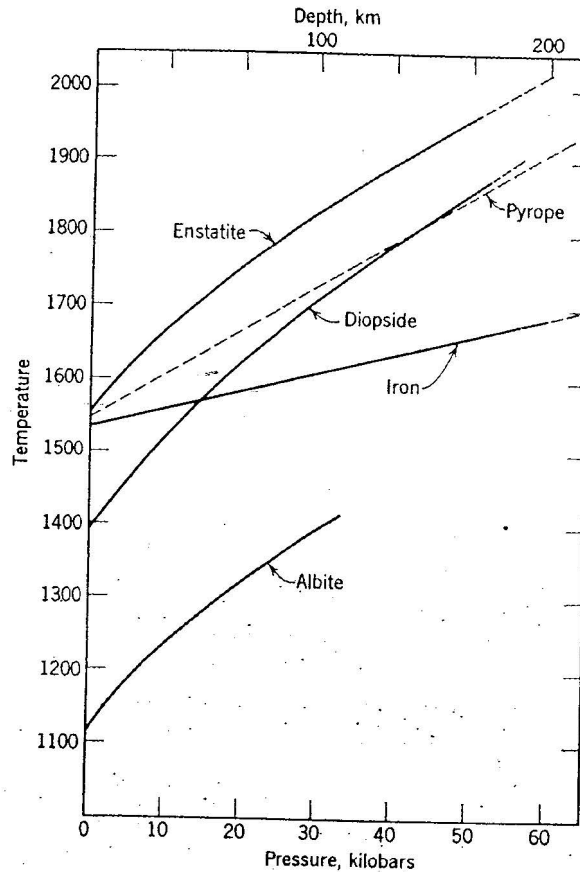
Figure 3.5 Melting temperatures versus pressure. (Boyd, 1964)
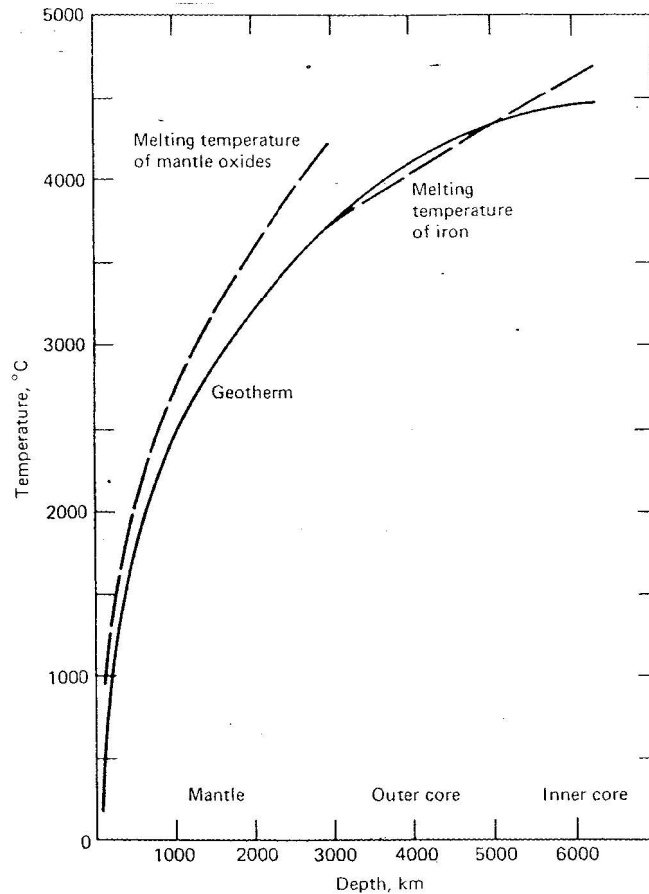
Figure 3.6   An estimated geotherm extending to the center of
the earth. Close approaches of mantle and core
melting temperatures to the geotherm take place
in the outer core and upper mantle.

THE INTERNAL STRUCTURE OF THE EARTH

The interpretation of seismic data provides a primary threefold division of the earth
into crust, mantle, and core (Figure 3.7). Together with other geophysical evidence,
these data also give some indication of the physical properties of the material making
up the three parts. It now remains to make plausible deductions as to the actual
constitution of the earth's interior, using both the above-mentioned evidence and
other significant information, such as the relative abundances of the elements and
the composition of meteorites.

The crust is directly accessible to our observations, in the upper part at least, and

general agreement exists as to its major features. It is heterogeneous and varies in thickness from place to place. Marked differences exist, especially between the continents and the deep ocean basins. The Mohorovičić discontinuity is found at different depths in different geological environments. Under all explored ocean basins it is between 10 and 13 km below sea level; under the continents it is usually about 35 km below sea level, descending to greater depths (as much as 60 km) below active mountain belts. The evidence indicates that in the ocean basins we have a depth of about 4 km of sea water, underlain by ½–2 km of unconsolidated sediments, followed by material of basaltic composition (5–8 km) to the Mohorovičić discontinuity.

It has been customary to interpret the seismic evidence on the continental crust as indicating two principal layers, an upper one of granite or granodioritic composition and a lower one of basaltic composition. These two layers correspond to the sial (i.e., material rich in Si and Al) and the sima (rich in Si and Mg), respectively, More refined seismic work, coupled with the geological heterogeneity of exposed continental areas, has largely refuted this concept of crustal layering. The continental crust is a mosaic of sediments, metamorphosed sediments, igneous intrusions of different kinds, and volcanics, faulted and broken into blocks of various shapes and sizes; however, there is probably a gradual change in average composition of the material, from
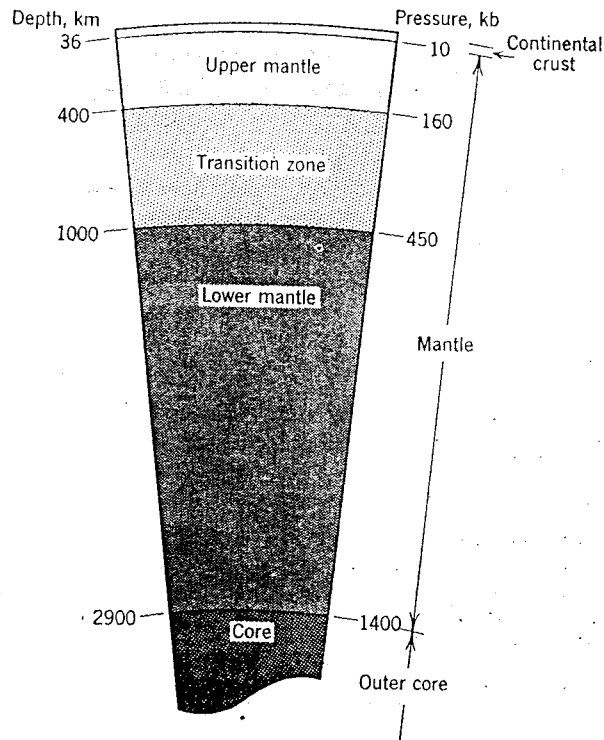
Figure 3.7 The internal structure of the earth.

granitic near the surface to gabbroic farther down. The earth's crust is in isostatic adjustment, the irregularities of the surface being compensated by the distribution of materials of different density within the crust and upper mantle.

The Mohorovičić discontinuity separates the heterogeneous crust from the more homogeneous mantle and is marked by a sudden increase in the velocities of seismic waves. The geophysical data indicate that the mantle has a layered structure; the upper mantle, to a depth of about 400 km, is separated from the lower mantle by a transition zone about 600 km thick. The nature of the upper mantle and the transition zone has been the subject of intensive study and extensive speculation in recent years. It is now realized that the key to many geological and geochemical problems—such as the origin of magmas, the triggering of deep-focus earthquakes, global tectonics, and continental drift—lies in the upper mantle and the transition zone. Our understanding of the nature of the upper mantle has been greatly increased by improved equipment and techniques for studying minerals and rocks at high temperatures and pressures. It is now possible to reproduce in the laboratory the physical conditions corresponding to depths of 400 km or more within the earth. The identification of the material making up the mantle is largely based on samples of possible mantle material brought up as inclusions in volcanic pipes, laboratory experiments on the behavior of minerals and rocks at high temperatures and pressures, and our knowledge of elemental abundances. Experiments indicate that only three rock types—dunite (olivine), peridotite (olivine and pyroxene), and eclogite (garnet and pyroxene)—have elastic properties of the right order to give the observed wave velocities in the upper mantle. This would also include material of chondritic meteorite composition, essentially a peridotite with a small content of plagioclase feldspar. These rocks are all made up largely of magnesium-iron silicates. Inclusions of dunite and peridotite are sometimes abundant in volcanic rocks (basalts and the diamond-bearing kimberlites); eclogite inclusions are less common. Basalt magmas on Hawaii appear to originate at depths of about 60 km, whereas the presence of diamonds in kimberlite indicates a greater depth of origin, over 100 km.

The nature of the Mohorovičić discontinuity has been the subject of considerable controversy. One school of thought considers it a physical discontinuity, the result of a phase change from lower crustal rocks of gabbroic composition to eclogite, which is higher-density material of essentially the same composition. The alternative view is that the discontinuity is a chemical one, the upper mantle having ultrabasic composition (dunite or peridotite). The problem has been reviewed by Ringwood (1975), who endeavored to correlate the geochemical, geophysical, and petrological information. He prefers an ultrabasic model for the upper mantle, with an overall composition corresponding to a mixture of one part basalt to three parts dunite, which he calls pyrolite (pyroxene-olivine rock). Fractional melting of this material would provide the basaltic magma, which has been copiously injected into and through the crust throughout geological time, and leave a residual dunite or peridotite. The mineralogy of pyrolite varies as a function of temperature and pressure, and Ringwood has shown that material of this composition could crystallize in four distinct assemblages, as follows:

(a) Olivine and amphibole   Ampholite

(b) Olivine + Al-poor pyroxene + plagioclase   Plagioclase pyrolite

(c) Olivine + Al-rich pyroxene + spinel   Pyroxene pyrolite

(d) Olivine + Al-poor pyroxene + garnet   Garnet pyrolite

Because the geotherms in the upper mantle differ considerably below continents and oceans, there are regional differences in the mineralogical composition of the pyrolite. The inferred composition of the upper mantle below Precambrian shield areas, average continental areas, and oceanic areas is illustrated in Figure 3.8. Under the Precambrian shields the mantle consists, to a considerable depth, of dunite or peridotite with minor segregations of eclogite. This layer is not as thick under the average continental areas and is thin or replaced by ampholite under the oceanic areas. Beneath the dunite-peridotite zone would be a lower density pyrolite zone. Evidence for such a region is found in the low-velocity seismic wave zone observed below 100 km.

An important development in our understanding of the transition zone in the mantle has been the realization of the significance of polymorphic changes taking place at high temperatures and pressures. This idea was originally due to Bernal, who in 1936 suggested that the discontinuity between the upper mantle and the transition zone results from the formation of a high-pressure polymorph of olivine. He pointed out that the analogous compound Mg_2GeO_4 exists in two modifications, one with the structure of olivine and the other an isometric phase with the spinel structure and 9% higher density.

Other polymorphic transformations that might be expected within the mantle are the change of $MgSiO_3$ from a pyroxene to an ilmenite structure, and of SiO_2 to a
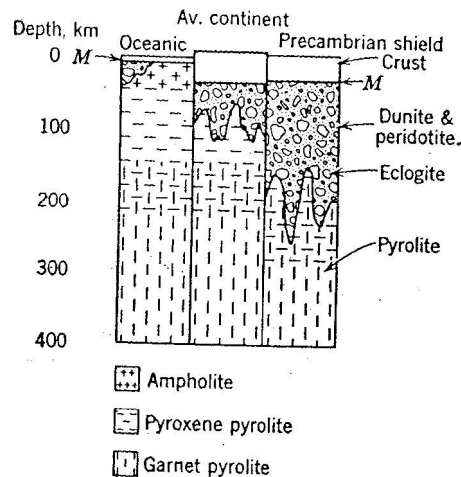
Figure 3.8 Petrological model for the upper mantle, as suggested by Clark and Ringwood (1964); M = Mohorovičić discontinuity.

rutile structure. Some of these predicted polymorphic changes have been confirmed by laboratory work. The rutile form of SiO_2 was made by Stishov and Popova in 1961 and later named stishovite; it is stable at about 130 kb and 1600°. Extrapolation of experimental results on $(Mg,Fe)_2SiO_4$ solid solutions shows that an olivine-spinel transition will occur in Mg_2SiO_4 at about 130 kb and 600°. Some substances with pyroxene structure are converted into ilmenite structures at high pressures, but laboratory experiments indicate that $MgSiO_3$ is probably converted into Mg_2SiO_4 and SiO_2 (stishovite) at about 120 kb in the temperature range 500°–2000°. All these transformations can be expected in the transition zone.

Clark and Ringwood have applied these data to the elucidation of the transition zone. As a first approximation they assume it is composed essentially of MgO and SiO_2, corresponding to a mixture of olivine, Mg_2SiO_4, and pyroxene, $MgSiO_3$. These phases would be stable to a depth of about 400 km. Below that depth a series of transformations into closer-packed phases would take place. In the light of available data, they postulate the following transformations with increasing depth:

$$2MgSiO_3 \text{ (pyroxene)} = Mg_2SiO_4 \text{ (olivine)} + SiO_2 \text{ (stishovite)}$$
$$Mg_2SiO_4 \text{ (olivine)} = Mg_2SiO_4 \text{ (spinel)}$$
$$Mg_2SiO_4 \text{ (spinel)} + SiO_2 \text{ (stishovite)} = 2MgSiO_3 \text{ (ilmenite)}$$
$$Mg_2SiO_4 \text{ (spinel)} = MgSiO_3 \text{ (ilmenite)} + MgO \text{ (periclase)}$$
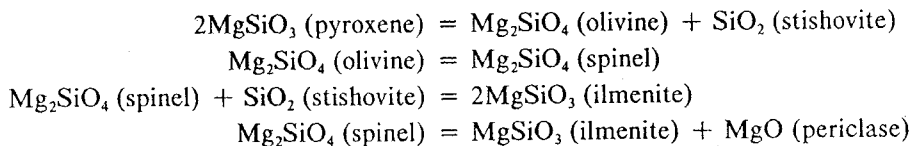
This series of transformations would become complete around a depth of 1000 km. They result in a density increase from 3.2 to 3.9 (referred to zero pressure). It is significant that the ultimate attainment of a close-packed state involves successive transformations through several intermediate states, reflected in the considerable depth of the transition zone. In addition, the mantle contains additional elements, mainly iron, calcium, aluminum, and sodium, which are in atomic substitution in the principal phases. Their presence influences the preceding reactions to some degree, in effect smearing out the transitions and resulting in a continuous rather than a stepwise increase in density.

The lower mantle, between 1000 and 2900 km, appears to be homogeneous, and presumably consists of a mixture of $(Mg, Fe)SiO_3$ with the ilmenite structure, and $(Mg, Fe)O$, periclase. The $FeO(FeO + MgO)$ ratio (molecular) has been estimated to be between 0.1 and 0.2. Ringwood has suggested the possibility of even more closely packed structures in the lower mantle, the ilmenite structure transforming to the perovskite $(CaTiO_3)$ structure and the periclase structure to the CsCl structure.

The belief that the earth has an iron core predates the seismic evidence for its existence. The idea apparently originated with Daubrée in 1866 and was based on the composition of meteorites. The concept of the iron core has become thoroughly entrenched in geophysical and geochemical thought, although it has been suggested that the core consists of the same material as the mantle but in a highly compressed form. However, the chemical difference between mantle and core has been confirmed by experiments with explosively generated shock pressures equal to those of the core. The results show that it is not possible to make a satisfactory core of light metals or their oxygen compounds. The physical properties of the core require elements of the

transition group, and only iron is sufficiently abundant. The properties of iron are close to those required and can be adjusted with small amounts of lighter elements. Geophysical evidence indicates that the mean atomic numbers of the core is about 22. Since the atomic number of iron is 26, this implies that the core contains elements of lower atomic number. Sulfur, carbon, and silicon have been suggested as possible elements. Sulfur and carbon, both relatively abundant in meteoritic matter, would require lower amounts than silicon to reduce the mean atomic number of the core.

The seismic discontinuity between the mantle and the core is a sharp one, and its position is known with considerable accuracy (± 2.5 km, according to Jeffreys). The passage of S waves through the core has not been observed (a few reported instances are generally regarded as spurious), and this is interpreted as indicating that the core is liquid at least in its outer part. Additional evidence for the fluidity of the core is obtained from the analysis of the bodily tide of the earth. Bullen suggested that the second-order discontinuity at a depth of about 5000 km is due to a change in rigidity and that below this depth the core is solid.

A plausible origin for the magnetic field of the earth and for its secular variation can be deduced from the presence of a liquid iron core, as shown by Elsasser. He suggests that the earth's magnetic field is the result of electric currents flowing in the interior and that such currents are to be expected in the highly conducting metal core rather than in the silicate mantle. The secular variation of the geomagnetic field is then interpreted as the expression of convection currents in the fluid core..

THE ZONAL STRUCTURE OF THE EARTH

→The earth may therefore be considered as made up of an iron core, a fairly homogeneous silicate mantle, and a heterogeneous silicate crust. This picture of its internal structure and composition is consistent with its mass, its moment of inertia, and the presence of discontinuities indicated by seismic waves. The decreasing silica content in the succession of materials from the crust into the mantle agrees with petrological experience. The probable composition of an earth of this kind is also consistent with the relative abundances of the elements. No one of these points, nor all of them together, proves that this picture is necessarily a true one, but at least it agrees with the available data. To complete the picture, we must add to the crust, mantle, and core three further zones: the atmosphere, the hydrosphere, and the biosphere. The atmosphere is the gaseous envelope that surrounds the earth. The hydrosphere is the discontinuous shell of water, fresh and salt, making up the oceans, lakes, and rivers. The biosphere is the totality of organic matter distributed through the hydrosphere, the atmosphere, and on the surface of the crust. Table 3.3 gives the important features of these zones, and Table 3.4 gives data on thickness, volume, mean density, and mass of each.

The atmosphere, the hydrosphere, and the biosphere, although geochemically important, contribute less than 0.03% of the total mass of the earth. Hence, in arriving at an average composition for the earth, these zones may be ignored. Even the crust makes up less than 1% of the whole. Thus the bulk composition of the earth is

Table 3.3 The Structure of the Earth

Name	Important Chemical Characters	Important Physical Characters
Atmosphere	N_2, O_2, H_2O, CO_2, inert gases	Gas
Biosphere	H_2O, organic substances, and skeletal matter	Solid and liquid, often colloidal
Hydrosphere	Salt and fresh water, snow, and ice	Liquid (in part solid)
Crust	Normal silicate rocks	Solid
Mantle	Silicate material, probably largely olivine and pyroxene or their high-pressure equivalents	Solid
Core or siderosphere	Iron-nickel alloy	Upper part liquid, lower part possibly solid

Table 3.4 Volume and Masses of Earth Shells*

	Thickness (km)	Volume ($\times 10^{27}$ cm^3)	Mean Density (g/cm^3)	Mass ($\times 10^{27}$ g)	Mass (percent)
Atmosphere	—	—	—	0.000005	0.00009
Hydrosphere	3.80 (mean)	0.00137	1.03	0.00141	0.024
Crust	17 "	0.008	2.8	0.024	0.4
Mantle	2883	0.899	4.5	4.016	67.2
Core	3471	0.175	11.0	1.936	32.4
Whole earth	6371	1.083	5.52	5.976	100.00

*Data for the biosphere is not included on account of its relatively small mass and the lack of precise figures.

essentially determined by that of the mantle and the core. However, in view of the significance of the crust in geology and geochemistry its composition is discussed in detail at this point.

THE COMPOSITION OF THE CRUST

The average composition of the crust is in effect that of igneous rocks, since the total amount of sedimentary and metamorphic rocks is insignificant in comparison to the bulk of the igneous rocks, and in any case their average composition is not greatly different. Clarke and Washington (1924) estimated that the upper 10 miles of the crust consist of 95% igneous and metamorphic rocks, 4% shale, 0.75% sandstone, and 0.25% limestone. Where sedimentary rocks are present they form a relatively thin veneer on an igneous basement, except where locally thickened in orogenic belts.

Clarke and Washington made an exhaustive study of the data available for computing an average composition of igneous rocks. The basis for their computations

was Washington's compilation of 5159 "superior" analyses. The analyses were grouped geographically, and the averages of these groups agreed fairly well with one another. In other words, the composition of of the earth's crust is approximately the same in different regions, provided these are large enough to eliminate local variations. However, the SiO_2 percentage is markedly lower for rocks from oceanic areas, such as the islands of the Atlantic and Pacific Oceans, further evidence for the belief that the sial shell is thin or absent below the ocean basins. Clarke and Washington's estimate of the average composition of igneous rocks, which is sometimes used as the average composition of the crust, is, in effect, the composition of continental areas rather than the crust as this term has been used earlier in this chapter.

The overall average of the 5159 analyses, recalculated to 100 with the elimination of H_2O and minor constituents, is as follows:

SiO_2	Al_2O_3	Fe_2O_3	FeO	MgO	CaO	Na_2O	K_2O	TiO_2	P_2O_5
60.18	15.61	3.14	3.88	3.56	5.17	3.91	3.19	1.06	0.30

This composition does not correspond to any common igneous rock but is intermediate between that of granite and basalt, which incidentally make up the bulk of all igneous rocks.

Numerous objections have been raised to the method of arriving at an average composition of igneous rocks by averaging analyses, but a more satisfactory procedure has yet to be proposed. The principal objections are

1. The uneven geographical distribution of analyses.
2. Their nonstatistical distribution over the different rock types.
3. The lack of allowance for the actual amounts of the rocks represented by the analyses.

The basis for the first objection is, of course, that Europe and North America have been more adequately investigated and are represented by far more analyses per unit-area than other parts of the earth. However, Clarke showed that the averages for the individual continental areas are in marked agreement in spite of the widely different coverage, a fact that suggests that the results can probably be accepted as a reasonable approximation.

The second objection is valid in that an average of published rock analyses must inevitably give undue weight to the rare and unusual rock types, and insufficient weight to the abundant and uniform types such as granites and basalts. This objection may not be as serious as has been asserted, since in a large number of analyses the unusual types will be drawn from the whole range of rock compositions and will tend to give a true average. Also, as Clarke and Washington point out, their 1924 average was made from analyses of fresh, unaltered rocks only, thus eliminating many analyses of unusual rocks.

The third objection, that all analyses are given equal weight regardless of the areas occupied by the rocks and therefore of their relative amounts, can be countered if it can be shown that in the average the variations due to this cause offset one another. It is true that one rock, say a basalt, is exceedingly abundant, whereas another may be merely a narrow dike. Such differences may not affect the mean appreciably, since

the relatively insignificant rocks range in composition from high silica to low silica just as the more abundant rocks do. Furthermore, the surface exposure of a rock is no certain measure of its real volume and mass, for a small exposure may be merely the peak or crest of a large subterranean body and a large exposure may represent only a thin layer.

An interesting confirmation of the general reliability of the figures of Clarke and Washington was provided by Goldschmidt. He suggested that if it were possible to obtain an average sample of a large part of the earth's crust consisting mainly of crystalline rocks, its analysis would give a reliable picture of the composition of the crust as a whole. Such an average sample was provided, he pointed out, by the glacial clay widely distributed in southern Norway. This clay represents the finest rock flour deposited by melt water from the Fennoscandian ice sheet. From 77 analyses of different samples of this material he calculated the following average analysis:

SiO_2	Al_2O_3	Fe_2O_3 + FeO	MgO	CaO	Na_2O	K_2O	H_2O	TiO_2	P_2O_5
59.12	15.82	6.99	3.30	3.07	2.05	3.93	3.02	0.79	0.22

These figures agree remarkably with those of Clarke and Washington, especially when the effects of hydration and solution, which result in the leaching of sodium and calcium, are taken into account.

Scientists at the Vernadsky Institute of Geochemistry in Moscow have made a systematic study of the composition of the crust, by analyzing many thousands of specimens and average samples of rocks from the Russian platform and the Caucasian geosyncline and combining these data with determinations from other regions. A comprehensive summary is provided by Ronov and Yaroshevsky (in Hart, 1969). They distinguish three types of crust—continental, oceanic, and subcontinental (essentially the transition zone of the continental shelf and slope)—each of which comprises shells of sedimentary and igneous rocks. Average thicknesses, masses, and mean compositions are derived for each of these shells, which are then combined to give averages for each of the three types of crust, and ultimately for the total crust (which they define as the material above the Mohorovičić discontinuity, rather than the 10-mile thickness used by Clarke and Washington). Their averages for continental crust and the total crust, calculated on a CO_2- and H_2O-free basis, are as follows:

	Continental Crust	Total Crust
SiO_2	61.9	59.3
TiO_2	0.8	0.9
Al_2O_3	15.6	15.8
Fe_2O_3	2.6	2.6
FeO	3.9	4.4
MnO	0.1	0.2
MgO	3.1	4.0
CaO	5.7	7.2
Na_2O	3.1	3.0
K_2O	2.9	2.4
P_2O_5	0.3	0.2

The second average reflects more truly than the figures of Clarke and Washington the composition of the crust as a whole, since it gives adequate consideration to the submarine region. However, the differences between the two sets of figures are small. Comparison with the average made by Clarke and Washington shows somewhat lower alkalies and higher magnesium and calcium, reflecting the basaltic nature of the submarine crust.

So far we have considered only the major elements, those that are determined as a matter of course in a rock analysis, and have omitted those present in lesser amounts. Obviously, the determination of the average amounts of the minor elements is a more difficult problem. However, in the last 40 years our knowledge of the relative and absolute abundances of the less common elements has greatly increased mainly as a result of the introduction of newer analytical methods such as quantitative spectrographic analysis, colorimetric determination, neutron activation, atomic absorption spectrophotometry, X-ray fluorescence, and isotope dilution. Three distinct procedures have been used in arriving at abundance figures for minor and trace elements: (a) the averages of many individual analyses; (b) the analysis of mixtures of many different rock types; and (c) the determination of the proportion of the trace element to some more common element with which it is geochemically associated (e.g., the abundance of rubidium can be estimated from the abundance of potassium and the average Rb:K). Table 3.5 presents the data on the average abundances of the elements in the earth's crust (illustrated graphically in Figure 3.9), and the abundances in a standard granite and a standard diabase. The data on the average crustal abundances are those for the continental crust (the suboceanic crust is probably close in composition to the average basalt); data for those elements more abundant than 1000 g/ton are taken from Clarke and Washington; for the minor and trace elements most of the figures are from a critical compilation by Taylor (1964), who combined averages for granites and basalts in the proportions of 1:1. The standard granite G-1 and diabase W-1 were originally prepared about 1948 from a granite from Westerly, Rhode Island, and a diabase from Centerville, Virginia, and have the distinction of being the most extensively analysed rocks in the world. These rocks were prepared to serve as geochemical standards; they were analyzed for major constituents by many laboratories in all parts of the world, and practically all the minor and trace elements have been determined in them by a wide variety of techniques. The figures given for the individual elements are considered the best choice from the available data, which not infrequently vary considerably. Even after over 25 years of analytical work changes are taking place in the selection of "best values." The data on these two rocks provide a meaningful comparison with those for the crustal average, based as that is on equal quantities of granite and basalt (the volcanic equivalent of diabase). Comparisons show that for many elements the figures are consistent throughout. For a few there are notable discrepancies, some of which can readily be explained—for example, G-1 is unusually high in thorium, a peculiar characteristic of the Westerly granite. Thanks in considerable part to the availability of G-1 and W-1, the quality and extent of minor and trace element data on rocks has improved markedly in recent years. It can fairly be said that the figures in Table 3.5 provide reliable norms for most elements, and thus make it possible to recognize abnormal

Table 3.5 The Average Amounts of the Elements in Crustal Rocks in Grams per Ton or Parts per Million*

Atomic Number	Element	Crustal Average	Granite (G-1)	Diabase (W-1)
1	H	1,400	400	600
3	Li	20	22	15
4	Be	2.8	3	0.8
5	B	10	1.7	15
6	C	200	200	100
7	N	20	59	52
8	O	466,000	485,000	449,000
9	F	625	700	250
11	Na	28,300	24,600	16,000
12	Mg	20,900	2,400	39,900
13	Al	81,300	74,300	79,400
14	Si	277,200	339,600	246,100
15	P	1,050	390	610
16	S	260	58	123
17	Cl	130	70	200
19	K	25,900	45,100	5,300
20	Ca	36,300	9,900	78,300
21	Sc	22	2.9	35
22	Ti	4,400	1,500	6,400
23	V	135	17	264
24	Cr	100	20	114
25	Mn	950	195	1,280
26	Fe	50,000	13,700	77,600
27	Co	25	2.4	47
28	Ni	75	1	76
29	Cu	55	13	110
30	Zn	70	45	86
31	Ga	15	20	16
32	Ge	1.5	1.1	1.4
33	As	1.8	0.5	1.9
34	Se	0.05	0.007	0.3
35	Br	2.5	0.4	0.4
37	Rb	90	220	21
38	Sr	375	250	190
39	Y	33	13	25
40	Zr	165	210	105
41	Nb	20	24	9.5
42	Mo	1.5	6.5	0.57
44	Ru	0.01		
45	Rh	0.005		<0.001
46	Pd	0.01	0.002	0.025
47	Ag	0.07	0.05	0.08
48	Cd	0.2	0.03	0.15
49	In	0.1	0.02	0.07
50	Sn	2	3.5	3.2
51	Sb	0.2	0.31	1.0
52	Te	0.01	<1	<1
53	I	0.5	<0.03	<0.03
55	Cs	3	1.5	0.9
56	Ba	425	1,220	160
57	La	30	101	9.8
58	Ce	60	170	23

Table 3.5   (continued)

Atomic Number	Element	Crustal Average	Granite (G-1)	Diabase (W-1)
59	Pr	8.2	19	3.4
60	Nd	28	55	15
62	Sm	6.0	8.3	3.6
63	Eu	1.2	1.3	1.1
64	Gd	5.4	5	4
65	Tb	0.9	0.54	0.65
66	Dy	3.0	2.4	4
67	Ho	1.2	0.35	0.69
68	Er	2.8	1.2	2.4
69	Tm	0.5	0.15	0.30
70	Yb	3.4	1.1	2.1
71	Lu	0.5	0.19	0.35
72	Hf	3	5.2	2.7
73	Ta	2	1.5	0.50
74	W	1.5	0.4	0.5
75	Re	0.001	<0.002	<0.002
76	Os	0.005	0.00007	0.0003
77	Ir	0.001	0.00001	0.003
78	Pt	0.01	0.0019	0.0012
79	Au	0.004	0.004	0.004
80	Hg	0.08	0.1	0.2
81	Tl	0.5	1.2	0.11
82	Pb	13	48	7.8
83	Bi	0.2	0.07	0.05
90	Th	7.2	50	2.4
92	U	1.8	3.4	0.58

*Omitting the rare gases and the short-lived radioactive elements.

enrichment or depletion of an element in any rock. The figures for G-1 and W-1 also show that some elements (such as Be, Rb, Ba) are markedly enriched in the granite, others (such as B, Sc, Ni) are enriched in the diabase, whereas some (such as Zn, Ga, Ge) show a rather uniform abundance in these contrasted rock types.

Some interesting features of Table 3.5 may be noted. Eight elements—O, Si, Al, Fe, Ca, Na, K, Mg—make up nearly 99% of the total. Of these oxygen is absolutely predominant. As Goldschmidt first pointed out, this predominance is even more marked when the figures are recalculated to atom percent and volume percent (Table 3.6). The earth's crust consists almost entirely of oxygen compounds, especially silicates of aluminum, calcium, magnesium, sodium, potassium, and iron. In terms of numbers of atoms oxygen exceeds 60%. If the volume of the different atoms, or rather ions, is calculated, oxygen makes up more than 90% of the total volume occupied by the elements. Thus the crust of the earth is essentially a packing of oxygen anions, bonded by silicon and the ions of the common metals. As Goldschmidt remarked, the lithosphere may well be called the oxysphere.

Table 3.5 shows that some of the elements that play a most important part in our
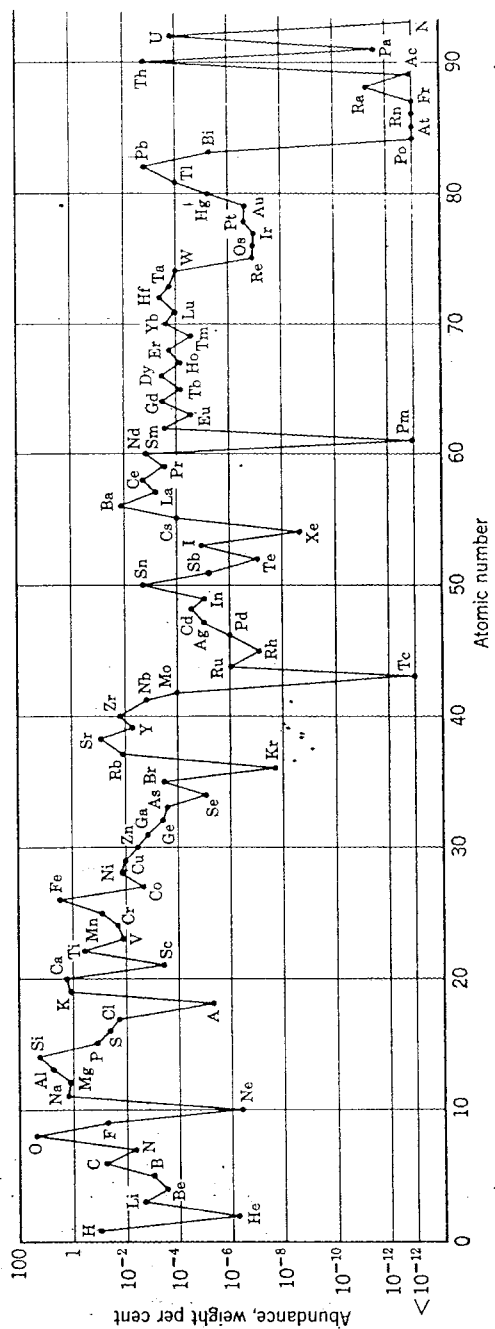
Figure 3.9 Crustal abundances of elements of atomic numbers 1 to 93.

Table 3.6   The Commoner Chemical Elements in the Earth's Crust

	Weight Percent	Atom Percent	Radius (Å)*	Volume Percent
O	46.60	62.55	1.40	91.7
Si	27.72	21.22	0.26	0.2
Al	8.13	6.47	0.53	0.5
Fe	5.00	1.92	0.77	0.5
Mg	2.09	1.84	0.72	0.4
Ca	3.63	1.94	1.12	1.5
Na	2.83	2.64	1.16	2.2
K	2.59	1.42	1.60	3.1

*The ionic radii used in this table and throughout the book are those listed in Appendix I.

economic life and that have long been known to and used by humans are actually quite rare. For example, copper is less abundant than zirconium; lead is comparable in abundance with gallium; mercury is rarer than any of the so-called "rare earths." Complementary to this is the relative abundance of many unfamiliar elements: rubidium is present in amounts comparable with nickel, vanadium is much more abundant than tin, scandium more abundant than arsenic, and hafnium, one of the last elements to be discovered, is more abundant than iodine. Evidently we must draw a clear distinction between the abundance of an element and its availability. Some elements, although present in the crust in considerable amounts, are systematically dispersed throughout common minerals and never occur in any concentration. Vernadsky called these *dispersed elements;* examples are rubidium, dispersed in potassium minerals, and gallium, in aluminum minerals. Other elements, such as titanium and zirconium, form specific minerals which, in turn, are widely dispersed in small amounts through some of the commonest rocks. The longest known and most familiar elements are those forming the major constituents of easily recognized minerals, minerals that are readily converted into useful industrial materials.

Fersman introduced the term *clarke,* defining it as the average percentage of an element in the earth's crust. Thus the clarke of oxygen is 46.60, of silicon, 27.72. In discussing dispersion and concentration of the elements, Vernadsky introduced a further term, the *clarke of concentration,* which is a factor showing the concentration of an element within a particular deposit or even a particular mineral. Thus, if the clarke of manganese is 0.1, the clarke of concentration of manganese in pyrolusite is 632, in rhodonite, 419, in a psilomelane with 50% Mn, 500. This factor is useful in the consideration of the migration and deposition of the elements and in the discussion of ore deposits. An ore is simply a deposit in which the concentration clarke of the element sought reaches a figure sufficient to make extraction profitable. Table 3.7 is an evaluation of the concentration clarkes necessary to form ore bodies of the commoner metals.

Table 3.7   Concentration Clarkes for Ore Bodies of the
Commoner Metals

Metal	Clarke	Minimum Percent Profitably Extracted	Concentration Clarke Necessary for an Ore Body
Al	8.13	30	4
Fe	5.00	30	6
Mn	0.10	35	350
Cr	0.01	30	3000
Cu	0.006	1	160
Ni	0.008	1.5	125
Zn	0.007	4	600
Sn	0.0002	1	5000
Pb	0.0013	4	3000
U	0.0002	0.1	500

The availability of an element depends in large part on its ability to form individual minerals in which it is a major constituent; the most unavailable elements are those that form no minerals of their own but occur in amounts generally much less than 1% in minerals of other elements—indium, rubidium, gallium, hafnium, rhenium, and so forth. Even for the "common" elements, however, the dispersed amounts exceed, by a vast factor, the amount available in so-called "ore deposits"; for example, clay minerals are a far more readily available source of aluminum than is bauxite, but the problem with the clay minerals is the technical one of extraction. Magnesium is extracted from sea water, in which it is present to the extent of 0.13%, although there are vast deposits of olivine that contain about 30%. The technical availability of a number of the rarer elements is conditioned in part by their inherently useful properties and in part by their being obtainable as byproducts from the extraction of more abundant elements. An example of the first kind is beryllium, which, although absolutely and relatively rare (its principal industrial source is as the mineral beryl, irregularly dispersed in a few granite pegmatites), has such valuable properties as an alloying element that it is an important industrial material. On the other hand, a number of exceedingly rare elements could be produced relatively easily as byproducts from extraction processes. For example, the electrolytic refining

Table 3.8   Annual World Consumption of the Elements (and their compounds) in Tons, in Powers of 10

10^9-10^{10}:	C
10^8-10^9:	Na, Fe
10^7-10^8:	N, O. S, K, Ca
10^6-10^7:	H, F, Mg, Al, P, Cl, Cr, Mn, Cu, Zn, Ba, Pb
10^5-10^6:	B, Ti, Ni, Zr, Sn
10^4-10^5:	Ar, Co, As, Mo, Sb, W, U
10^3-10^4:	Li, V. Se, Sr, Nb, Ag, Cd, I, Rare Earths, Au, Hg, Bi
10^2-10^3:	He, Be, Te, Ta

of copper can provide tellurium in sufficient quantities to encourage the development of industrial uses of this element. Similarly, Goldschmidt pointed out that if any demand for gallium or germanium existed large amounts of these elements (1000 tons or so annually) could be extracted from the ashes of certain coals. The demand for germanium for industrial purposes has resulted in the application of Goldschmidt's suggestion to the production of commercial amounts of this element. The annual world consumption of the elements and their compounds is summarized in Table 3.8.

THE COMPOSITION OF THE EARTH AS A WHOLE

The bulk composition of the earth is essentially determined by the composition and relative amounts of the mantle and the core, since they make up over 99% of its mass. By using the deductions regarding the earth's interior discussed above, an average composition for the earth as a whole can be calculated.

For the purpose of this calculation the following assumptions about the composition of the core and the mantle are made:

1. The iron alloy of the core has the composition of the average for nickel-iron in chondrites, and includes the average amount (5.3%) of FeS in these meteorites.
2. The composition of mantle plus crust is the same as the oxidic material (silicates plus small amounts of phosphates and oxides) of the average chondrite.

These assumptions are certainly oversimplifications. For example, all Ni is assumed to be in the core, whereas terrestrial olivine always contains nickel, that from peridotite inclusions in basalt usually having 0.2–0.3% Ni. Some nickel is certainly combined in mantle silicates. Again, the assumption that all iron sulfide is incorporated in the core is clearly an approximation at best, for some sulfide is certainly present in mantle and crust. However, the incorporation of iron sulfide in the core rather neatly meets the geophysical requirement that the density of the outer core is less than that of iron by about 10%. If 27.1% nickel-iron, density 7.90, is mixed with 5.3% troilite, density 4.80, the resulting mixture will have a density, at zero pressure, of 7.15, thereby meeting this requirement quite closely. This obviates the necessity of incorporating Si in the core, as has been suggested by Ringwood and others. A nickel-iron core with free Si is also incompatible with a mantle containing appreciable amounts of FeO, if mantle and core are in chemical equilibrium.

Mass of core equals 32.4% mass of earth and mass of mantle plus crust equals 67.6% mass of earth; the core is assumed to consist of nickel-iron of composition of average chondrite plus 5.3% troilite, and mantle and crust are assumed to have a composition of silicate material of average chondrite.

The results of this calculation are given in Table 3.9. Ganapathy and Anders (1974) have calculated the bulk composition of the earth on a rather more elaborate set of assumptions and their figures are included for comparison.

In view of the uncertainties and assumptions involved in calculations of this kind, too much weight should not be given to the numerical results obtained. Nevertheless, some significant deductions of a semiquantitative order are probably justified. The

results indicate that about 90% of the earth is made up of four elements, Fe, O, Si, and Mg. The only other elements that may be present in amounts greater than 1% are Ni, Ca, Al, and S. Seven elements, Na, K, Cr, Co, P, Mn, and Ti may occur in amounts from 0.01 to 1%. Thus the earth is made up almost entirely of 15 elements, and the percentage of all the others is negligible, probably 0.1% or less of the whole.

Let us now compare the order of abundance of the major elements in the earth with their relative abundances in other parts of the solar system. Table 3.10 gives these data for the earth as a whole, for the earth's crust, for the average composition of meteorites, and for the sun and moon. The most striking feature is the relative uniformity throughout; the same elements appear in all columns, albeit in different order. The rarity of hydrogen and helium in the earth and meteorites is, of course, easily understood. When we come to consider the nonvolatile elements, it is seen that iron, silicon, and magnesium in general head the lists and are followed by nickel, sodium, calcium, and aluminum. Note particularly that the most abundant elements are all of low atomic number; none with atomic number greater than 30 appears in any of the lists, except for Ba in the crust.

The relative abundances of the elements in the crust are included in Table 3.10 for comparison with the bulk composition of the earth. The main differences are the lesser abundance of iron and magnesium in the crust, the nonappearance of nickel and sulfur, and the increased significance of aluminum, potassium, and sodium. This suggests that the differentiation of the earth has led to a concentration of relatively light, easily fusible alkali-aluminum silicates at the surface.

A comparison of the elemental composition of the mantle plus crust (the silicate material of Table 3.9) with that of the crust (Table 3.5) reveals some interesting

Table 3.9   Calculation of the Bulk Composition of the Earth

	Metal	Troilite	Silicate	Total	Total*
Fe	24.58	3.37	6.68	34.63	35.98
Ni	2.39			2.39	2.02
Co	0.13			0.13	0.093
S		1.93		1.93	1.66
O			29.53	29.53	28.65
Si			15.20	15.20	14.76
Mg			12.70	12.70	13.56
Ca			1.13	1.13	1.67
Al			1.09	1.09	1.32
Na			0.57	0.57	0.143
Cr			0.26	0.26	0.472
Mn			0.22	0.22	0.053
P			0.10	0.10	0.213
K			0.07	0.07	0.017
Ti			0.05	0.05	0.077
	27.10	5.30	67.60	100.00	100.688

*Ganapathy and Anders, *Lunar Science*, V, p. 256, 1974.

Table 3.10 Relative Abundances (by Weight) of the Elements

Crust*	Whole Earth†	Meteorites‡	Sun§	Moon‖
O	Fe	O	H	O
Si	O	Fe	He	Si
Al	Si	Si	O	Mg
Fe	Mg	Mg	C	Fe
Ca	Ni	S	N	Ca
Na	S	Ni	Si	Al
K	Ca	Ca	Mg	Ni
Mg	Al	Al	Fe	S
Ti	Na	Na	S	Ti
H	Cr	Cr	Al	Cr
P	Mn	Mn	Ca	Na
Mn	Co	P	Ni	P
F	P	Co	Na	Mn
Ba	K	K	Cr	V

*From Table 3.5.
†From Table 3.9.
‡From Table 2.4.
§From Table 2.3.
‖ Ganapathy and Anders, *Lunar Science*, V, p. 256, 1974.

fractionations. The mass of the crust (the solid material above the Mohorovičić discontinuity) is 0.024×10^{24} g, that of the mantle 4.016×10^{27} g, so the mass of the crust is only 0.6% that of the mantle. The major elements O and Si and the minor elements Mn and P show no marked fractionation between mantle and crust. In the crust relative to the mantle, Mg is much diminished; Cr is also strongly diminished, with the average in the crust about 0.02%; Fe, on the other hand, is only slightly less concentrated in the crust than in the mantle. Elements Al and K are strongly enriched in the crust, Na and Ca to a lesser extent—in terms of mineralogical composition the differentiation of crust from mantle has largely been the accumulation of feldspar in the crust.

THE PRIMARY DIFFERENTIATION OF THE ELEMENTS

It seems probable that the earth was formed by the gradual accretion of solid planetesimals—indeed, the chondritic meteorites may be parts of planetesimals left over from the preplanetary stage of the solar system. The chondritic meteorites are made up of three different phases or groups of phases: nickel-iron, iron sulfide, and silicate minerals (mainly olivine and pyroxene). The elements are distributed between these phases according to their relative affinity for metal, for sulfide, or for silicate. The composition of these phases in meteorites has been essentially determined by equilibrium in the Fe—Mg—Si—O—S system, in which oxygen greatly exceeded sulfur and the sum of the two was insufficient to combine completely with the electropositive elements. Since Fe was more abundant than Mg and Si, most easily reduced to

the metal, and with the greatest affinity for sulfur, a system of three essentially immiscible phases—iron-magnesium silicate, iron sulfide, and free iron—resulted, and the amount of combined iron was fixed by the amount of oxygen plus sulfur. The distribution of the remaining electropositive elements was governed by reactions of the type

$$M + Fe \text{ silicate} \rightleftharpoons M \text{ silicate} + Fe$$
$$M + Fe \text{ sulfide} \rightleftharpoons M \text{ sulfide} + Fe$$
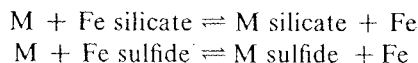
that is, by the free energies of the corresponding silicates and sulfides in relation to those of iron sulfide and iron silicate.

In broad outline, we can perceive the factors that determine the distribution of the elements in a system of this kind. Iron, because of its preponderant abundance, was common to all the condensed phases—the metallic phase rich in free electrons, the ionic silicate phase, and the semimetallic, probably covalent, sulfide phase. Elements more electropositive than iron displaced iron from the silicates, in which they were accordingly concentrated. Conversely, less electropositive elements concentrated in the metal, being displaced by iron from ionic compounds. The sulfide phase attracted those elements that form essentially homopolar compounds with sulfur and the metalloids and that cannot exist in an ionic environment with appreciable concentrations of metalloid ions; such elements are in the main the metals of the sulfide group of analytical chemistry.

It is highly significant that the distribution of elements in a gravitational field, such as that of the earth, is controlled not by their densities or atomic weights, as might perhaps be expected, but by their affinities for the major phases that can be formed. This is, in turn, controlled by the electronic configurations, hence the chemical bonding characteristics, of their atoms. For example, uranium and thorium, although of high density, are strongly electropositive elements and have concentrated in the earth's crust as oxides or silicates. Gold and platinum metals, on the other hand, have no tendency to form oxides or silicates, and alloy readily with iron; hence they are presumably concentrated in the earth's core. The distribution of the elements is not directly controlled by gravity, which merely controls the relative positions of the phases; the distribution of the elements within these phases depends on chemical potentials.

THE GEOCHEMICAL CLASSIFICATION OF THE ELEMENTS

Goldschmidt was the first to point out the importance of this primary geochemical differentiation of the elements. He coined the terms siderophile, chalcophile, lithophile, and atmophile to describe elements with affinity for metallic iron, for sulfide, for silicate, and for the atmosphere, respectively. When he put forward this concept in 1923, few quantiative data were available on which to base his ideas. The geochemical nature of an element could, of course, be established by measuring its distribution between three liquid phases of metal, sulfide, and silicate. He recognized the difficulty of carrying out these measurements in the laboratory, but remarked

that meteorites provide us with such an experiment in a fossilized condition. Many meteorites consist of nickel-iron, troilite, and silicate, all of which have presumably solidified from a liquid state. The distribution of a particular element between those three phases would be established when the system was in a liquid condition and can be determined by mechanically separating these phases and analyzing them individually. From such analyses the partition of the element between metal, silicate, and sulfide can easily be calculated.

In later years Goldschmidt and his co-workers made many measurements of the content of various elements in nickel-iron, troilite, and silicate of meteorites. Much work along these lines was also carried out by I. and W. Noddack in Berlin. These measurements have been greatly refined and extended in recent years by the application of new techniques, especially neutron activation analysis. The information derived from meteorites was supplemented by that obtained by a study of smelting processes, such as the distribution of elements during the smelting of the Mansfeld copper slate in Germany. The smelting of this slate gives a silicate slag, a matte rich in iron and copper sulfide, and metallic iron. Spectrographic measurements of the concentration of many minor elements in the different phases have given distribution coefficients agreeing on the whole with those determined from examination of meteorites.

On the basis of these results an element may be classified according to its geochemical affinity into one of four groups: siderophile, chalcophile, lithophile, and atmophile (Table 3.11). Some elements show affinity for more than one group, because the distribution of any element is dependent to some extent on temperature, pressure, and the chemical environment of the system as a whole. For instance, chromium is a strongly lithophile element in the earth's crust, but if oxygen is deficient, as in some meteorites, chromium is decidedly chalcophile, entering almost exclusively into the sulfo-spinel daubréelite, $FeCr_2S_4$. Similarly, under strongly reducing conditions carbon and phosphorus are siderophile. The mineralogy of an element,

Table 3.11   Geochemical Classification of the Elements (based on distribution in meteorites)

Siderophile	Chalcophile	Lithophile	Atmophile
Fe* Co* Ni*	(Cu) Ag	Li Na K Rb Cs	(H) N (O)
Ru Rh Pd	Zn Cd Hg	Be Mg Ca Sr Ba	He Ne Ar Kr Xe
Os Ir Pt	Ga In Tl	B Al Sc Y La-Lu	
Au Re† Mo†	(Ge) (Sn) Pb	Si Ti Zr Hf Th	
Ge* Sn* W‡	(As) (Sb) Bi	P V Nb Ta	
C‡ Cu* Ga*	S Se Te	O Cr U	
Ge* As† Sb†	(Fe) Mo (Os)	H F Cl Br I	
	(Ru) (Rh) (Pd)	(Fe) Mn (Zn) (Ga)	

*Chalcophile and lithophile in the earth's crust.
†Chalcophile in the earth's crust.
‡Lithophile in the earth's crust.

Table 3.12   The Geochemical Classification of the Elements in Relation to the Periodic System

Legend (center of table):
Atmophile: N
Lithophile: Na
Chalcophile: Zn
Siderophile Fe

H																	He
Li	Be											B	C	N	O	F	Ne
Na	Mg											Al	Si	P	S	Cl	Ar
K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	Kr
Rb	Sr	Y	Zr	Nb	Mo		Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I	Xe
Cs	Ba	La-Lu	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi			
			Th	U													

although a general guide, may not be altogether indicative of its geochemical character. For example, although all thallium minerals are sulfides, the greater part of the thallium in the earth's crust is contained in potassium minerals, in which the Tl^+ ion proxies for the K^+ ion. In general, the classification of an element as lithophile, chalcophile, or siderophile refers to its behavior in liquid-liquid equilibria between melts. When an element shows affinity for more than one group, it is given in parentheses under the group or groups of secondary affinity.

The geochemical character of an element is largely governed by the electronic configuration of its atoms and hence is closely related to its systematic position in the periodic table (Table 3.12). Lithophile elements are those that readily form ions with an outermost 8-electron shell; the chalcophile elements are those of the B subgroups, whose ions have 18 electrons in the outer shells; the siderophile elements are those of Group VIII and some neighboring elements, whose outermost shells of electrons are for the most part incompletely filled. These factors are reflected by other properties also. Goldschmidt pointed out the marked correlation between geochemical character and atomic volume. If the atomic volume of the elements is plotted against atomic number, the resulting curve shows maxima and minima. All siderophile elements are near the minima; the chalcophile elements are on sections in which the atomic volume increases with the atomic number; they are followed by the atmophile elements, whereas the lithophile elements are near the maxima and on the declining sections of the curve.

If the heat of formation of an oxide is greater than that of FeO the element is lithophile; the difference between the two heats of formation is a measure of the intensity of the lithophile character. Similarly, those elements having oxides with heats of formation lower than FeO are chalcophile or siderophile. Semiquantitative measure of lithophile, siderophile, or chalcophile character is also provided by the electrode potential. Elements with high positive potentials (1–3 volts), such as the alkali and alkaline earth metals, are lithophile; the noble metals, with high negative potentials, are siderophile; elements falling in the intermediate range are generally chalcophile.

THE PREGEOLOGICAL HISTORY OF THE EARTH

The pregeological history of the earth comprises the sequence of events through which it passed before the time when the physical condition of the surface became much as it is today—a surface partly of rocks, partly of water, with a mean temperature essentially determined by solar radiation. This is the zero datum for geological time, a datum following which the earth's surface has been subject to the normal processes of weathering and erosion. The pregeological history began when the earth originated as an individual body within the universe. We do not know the length of time covered by this period with any degree of precision; the earth's crust has existed for over 3700 million years, and the earth as an individual planetary body for 4600 million years.

As discussed in Chapter 2, the earth is believed to have formed from the same material that gave rise to the sun and the other planets. Hypotheses differ as to the mode of aggregation; one considers that the earth condensed from incandescent gas, the other that it grew by the gradual accretion of solid particles in a cosmic dust cloud. With these things in mind it is interesting to compare the composition of the earth with that of the sun. The major differences are the great abundance of hydrogen and helium and to a lesser extent carbon and nitrogen in the sun. These differences can be adequately explained on either hypothesis, since these elements are either gaseous or form stable gaseous compounds. If the earth grew by the accretion of solid particles, it is readily seen that these elements have always been of low abundance on the earth; if the earth was formed by the condensation of a mass of incandescent material, the light gases would have tended to escape from the earth's gravitational field during cooling from the high temperatures. One piece of evidence, however, favors the accretion hypothesis; not only are the light gases of low abundance on the earth but the heavy gases are also. Krypton and xenon are about a million times less abundant than their immediate neighbors in the periodic table, a definite deficiency compared to their probable abundance throughout the universe. If the earth condensed from incandescent material, these gases should be more abundant than they are; if the earth was formed by the accretion of solid particles, then they were never present in any amount. On the accretion hypothesis the earth's atmosphere was formed by the release of gases occluded and chemically combined in the solid particles.

However, although the earth and the other planets probably aggregated from the same material that gave rise to the sun, their individual compositions differ considerably. These differences can be ascribed at least in part to the extreme temperature differential between the center and the margin of the solar system, whereby the outer planets are mantled by great thicknesses of condensed gases (H_2O, NH_3, CH_4, etc). Differences in density between the inner planets can be explained by variations in the relative proportions of combined and free iron for similar bulk compositions. Possibly, however, considerable differences in bulk composition exist even between the inner planets, such as certainly exist between the earth and the moon. Such differences may have arisen from chemical fractionations in the ancestral solar nebula, and from nonhomogeneous accretion processes, whereby the composition of the

Table 3.13   Three Possible Models for the Formation of the Earth

Planetary Disk of Solar Composition		
I	II	III
Formation of planetisimals of approximate ordinary chondrite composition (metallic iron and silicate)	Formation of fine-grained material resembling carbonaceous chondrites	Formation of compounds of increasing volatility with decrease in temperature
↓	↓	↓
Accumulation	Accumulation	Accumulation on rapid time scale of high-temperature fraction (mainly metallic iron and silicates)
		↓
Segregation into core, mantle, crust, atmosphere, and hydrosphere, possibly over long time scale	Autoreduction	Segregation into core and mantle
	↓	↓
	Segregation of core and release of CO to atmosphere	Veneer segregated planet with low-temperature condensates, include water as hydrated silicates
	↓	↓
	Loss of primitive atmosphere	Process crust and upper mantle to form present configuration
	↓	
	Formation of crust, atmosphere, and hydrosphere soon after from residual	

Turekian, *Chemistry of the Earth,* p. 99, 1972. Courtesy of Holt, Rinehart and Winston, Inc.

accreting material changed during accretion. Turekian has outlined three possible models (Table 3.13) for the formation of the earth. His Model III typifies a nonhomogeneous accretion process, whereby a crust of markedly different bulk composition is added to the earlier accreted core and mantle.

Whichever theory of origin we accept, the internal structure of the earth with its marked density stratification seems to demand that at an early period in its history it was sufficiently hot for metallic iron to liquify and undergo gravitational accumulation to form the core, If, as appears probable, the earth formed by the accretion of planetesimals resembling chondritic meteorites, the initial state of our planet was an intimate mixture of nickel-iron, troilite, and silicate minerals, broadly homogeneous throughout. The time of formation of the earth by this accretion of planetesimals is usually placed 4.6×10^9 years ago, the figure obtained for the time when certain stony meteorites were enriched in and certain iron meteorites were depleted in uranium and thorium. However, it is not an obvious fact that the latter event and

the formation of the earth were necessarily contemporaneous, and there has been a tendency to push back the time of accretion of the earth to about 5×10^9 years. The identification in the Allende carbonaceous chondrite, of high-temperature inclusions which include oxides, silicates and titanates, has led some investigators to consider models in which these refractory materials were the first materials to condense in the solar system. Such a sequence would require modifications in the heterogeneous formation model given in Table 3.11. The condensation of refractory oxides has also been suggested to account for the chemistry of the moon, including its low concentrations of volatile materials.

The next stage after accretion of the earth was one of heating, mainly by radioactivity. Here it is essential to bear in mind the greater effect of the shorter-lived radioactive nuclides as we go back in time. This is graphically illustrated in Figure 3.10, which is based on the known heat production by radioactive nuclides and their estimated abundances in the earth. This figure shows that at 4.5×10^9 years ago radiogenic heat production in the earth would be at least seven times greater than today, and most of the heat was contributed by K^{40}. In addition, even shorter half-lived nuclides such as Al^{26} may have contributed significant amounts of heat energy.

Under these circumstances the roughly homogeneous undifferentiated earth would heat up comparatively rapidly. With the chondritic values for the radioactive elements, the temperatures within the primitive earth would increase as shown in Figure 3.11. On this basis, the melting temperature of iron was reached in about 600 million years at a depth of a few hundred kilometers; at this stage a change in regime sets in, and the curves for later times are not realized.

Elsasser (1963) has cogently linked this thermal evolution with the formation of the iron core. He points out that the thermal evolution depicted in Figure 3.11 results in the melting of the free iron in a zone at a depth of several hundred kilometers and the accumulation of this metal into a layer of molten iron. A coherent layer of molten iron within a predominantly silicate mantle is clearly unstable, because material of higher density overlies that of lower density. Elsasser shows that this results in the development of a large "drop" (Figure 3.12), which then sinks toward the center of the earth, displacing the lighter silicates. It is essentially the mechanism of salt dome or igneous intrusion, on a giant scale and in the reverse direction. The sinking of this large mass of molten iron to the center would transform a considerable amount of gravitational energy into heat, sufficient to raise the temperature of the interior of the earth by some 2000°. Thus the process of core formation was strongly exothermic and self-accelerated, and Elsasser believes that the time required for such a "drop" to form and fall was of the order of 100,000 years.

With the movement of the free iron, approximately one-third of the total mass of the earth, toward the center, the whole planet was profoundly reorganized. While the iron was unsymmetrically sinking the lighter fractions of the silicate material were unsymmetrically rising, undergoing partial fusion, reaction of solid material with melt, followed by fractional crystallization. Under these circumstances it is not surprising that crustal rocks older than 4×10^9 years have not been found. The development of thick, stable crustal blocks had to await the decay of radioactivity and the solidification of most of the mantle.
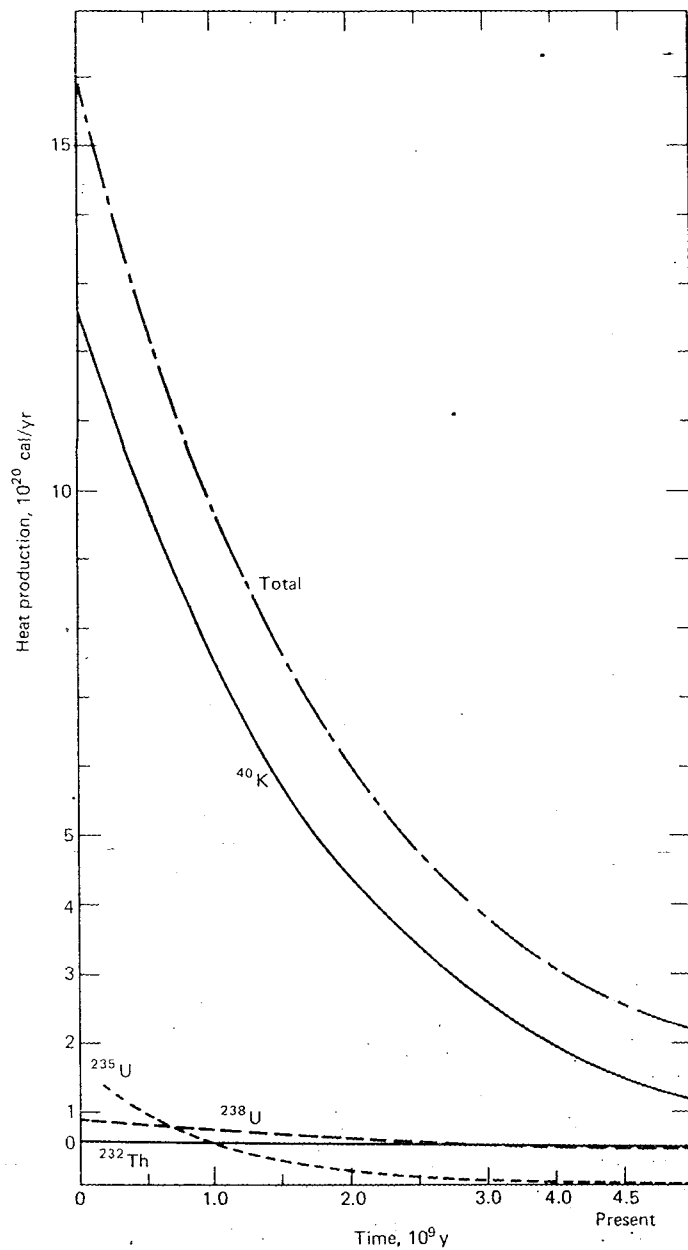
Figure 3.10 Radiogenic heat formed in the Earth by U, Th and ^{40}K separately and together plotted against time.
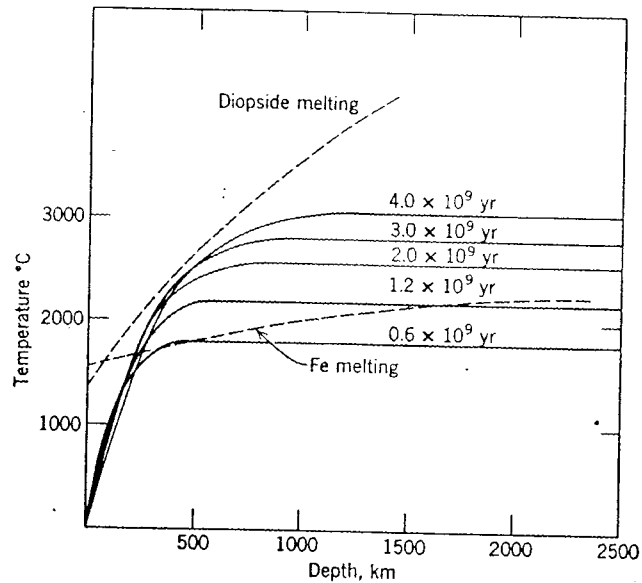
Figure 3.11 Temperatures in a homogeneous Earth heated by the radioactivity of the average chondrite. (MacDonald)

The development of the crust is a further problem and one of special importance to geochemistry. Considerable diversity of opinion exists. One school of thought pictures the entire crust of the earth solidifying as a basalt. No differentiation into continents and ocean basins would then exist. Localized fractures in this initially thin basaltic crust would channel lava, hot gases, and solutions from the hotter subcrustal layers, and these fractures would provide nuclei (something like present-day island arcs) for the nascent continents. The action of solutions carrying silica and alkalies would produce less basic crustal rocks. Erosion and sedimentation, at first on a small scale, would accelerate the chemical differentiation of crustal material. This concept pictures the continents growing through pregeological (and geological) time by the successive welding of sedimentary deposits and new island arcs to the initial continental nuclei.

The other school of thought pictures the initial continents as the final product of the solidification of the crust. It carries the differentiation by crystallization, already exemplified by the separation of basaltic crust from peridotitic mantle, a step farther
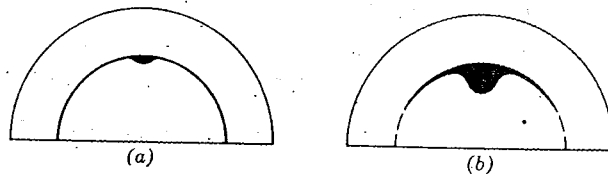
Figure 3.12 Stages of drop formation from a heavy liquid layer. (Elsasser, 1963)

and considers the continental nuclei as being raft-like masses essentially of quartz and feldspar (i.e., of granitic or granodioritic composition). An asymmetrical distribution of continents and oceans may have been produced by the asymmetrical production of the large iron "drops." The first protocontinent would form in the region above the first large falling drop, and the antipodal region would remain oceanic.

Present evidence suggests that the development of a stable crust was not a rapid, once-for-all process. The interval between the age of the earth and the age of the oldest dated rocks is significant. Great crustal instability during this interval is indicated. The original continental nuclei may have been partly remelted and rebuilt many times before they grew to a sufficient size and thickness to resist further engulfment. However, by about 3.5×10^9 years ago the pattern of continental nuclei and ocean basins was probably established, since rocks near this age are known in Greenland, Minnesota, and South Africa.

The origin of the atmosphere and hydrosphere is discussed in detail in later chapters. The growth of the earth by accretion in a solar dust cloud suggests an initial atomosphere lacking free oxygen—it may have consisted largely of hydrogen, helium, nitrogen, water vapor, methane, and ammonia (methane and ammonia are present in the atmospheres of the larger planets) or as carbon dioxide and nitrogen if hydrogen was low in abundance. The gravitational attraction of the earth is insufficient to retain hydrogen and helium, and any present would gradually diffuse into outer space. The evolution of the primeval atmosphere into the present one, which consists essentially of nitrogen and oxygen, is pictured as beginning with the photochemical dissociation of water vapor by solar radiation, thereby producing free oxygen. Any methane would then oxidize to carbon dioxide and water. Other photochemical reactions would produce more complex organic compounds, ultimately leading to self-reproductory systems, in effect living matter. The possibility of such reactions has been demonstrated in the laboratory by Miller (1957), who passed electrical discharges through a mixture of methane, ammonia, water vapor, and hydrogen and was able to demonstrate the presence of several carboxylic and amino acids in the products.

The various suggestions as to how life originated are summarized by Cloud (1976). The initial development of living matter was dependent on the preexistence of organic molecules such as amino acids. However, once cells of a kind were established photosynthesis became possible. Photosynthesis was a revolutionary event in earth history; it enabled organisms to collect and use solar radiation directly and thereby utilize carbon dioxide to synthesize more complex compounds; it was also responsible for the gradual release of oxygen into the atmosphere. Photosynthesis may have originated through the development of certain colored compounds capable of acting as catalysts in photochemical reactions. These compounds were probably porphyrins, which are easily synthesized and are extremely stable for organic substances. Porphyrins are formed by the condensation of four pyrrole derivatives around a metal atom [pyrrole is a five-membered ring compound with the formula $(CH)_4NH$]. It is extremely suggestive that porphyrins act as light receivers in the chlorophyll of plants, as oxygen carriers in many animals, and as respiratory pigments in cells. Once photosynthesis was established, a vastly greater supply of energy

was available for organisms, whether they used light themselves or ate others that did so.

The length of the period between the establishment of a stable crust and the appearance of life is difficult to evaluate. The oldest well-dated fossils are the microflora in the Gunflint chert, a Precambrian formation on the north shore of Lake Superior, whose age is fairly well established at 1.9×10^9 years. Stromatolites, laminated calcareous structures interpreted as fossil algae, are known from the Bulawayan System of Zimbabwe, believed to be older than 2.6×10^9 years, but neither the interpretation nor the age is firmly established. Most of the oldest sedimentary rocks are conglomerates, sandstones, and shales, showing few features indicating that conditions in the atmosphere and hydrosphere were very different from those today. However, Cloud (1976) and others have pointed out a number of suggestive features. The oldest formations show a notable dearth of limestones, which first become abundant in the younger Precambrian deposits, with ages less than 1.2×10^9 years; this indicates a rather high partial pressure of CO_2 in the atmosphere and hydrosphere, acting to keep $CaCO_3$ in solution. Red beds, colored by ferric oxide produced by oxidative weathering, appear about the same time, suggesting an increase in atmospheric oxygen over that present earlier. The most characteristic of Precambrian sediments, which are apparently unique to this system and are not repeated in later times, are banded-iron formations of the Lake Superior type, none of which is younger than about 1.7×10^9 years. These banded-iron formations occur on all the continents, and their formation implies a nonoxidizing, CO_2-rich atmosphere under which large quantities of iron could be transported to the sites of deposition in the soluble ferrous state. The combined evidence of geochemistry, paleontology, and stratigraphy indicates that photosynthesis by green plants was probably established at least 2.0×10^9 years ago, but that atmospheric oxygen was first available in relatively large quantities about 1.2×10^9 years ago. By the end of the Precambrian organic evolution had produced the metazoa (multicelled organisms), and probably by that time the atmosphere and hydrosphere were not notably different geochemically from their present state.

SELECTED REFERENCES

Anders, E. (1977). Chemical compositions of the moon, earth, and eucrite parent body. *Phil. Trans. R. Soc. Lond. A* **285**, 23–40. A systematic approach to the estimation of planetary compositions.

Anders, E. and T. Owen (1977). Mars and earth: origin and abundance of volatiles. *Science*, **198**, 453–465. An important paper giving evidence for the accretion of planetary bodies with differing amounts of different primary components.

Birch F. (1965). Speculations on the earth's thermal history. *Bull. Geol. Soc. Amer.* **76**, 133–154. A critical examination of the development of core, mantle, and crust as a result of radioactive heating in an originally cool accreted body.

Boyd, F. R. (1964). Geological aspects of high-pressure research. *Science* **145**, 13–20. An informative review article on the application of laboratory investigations to the nature of the earth's interior.

Bullen, K. E. (1963). *An introduction to the theory of seismology* (fourth ed.). 381 pp. Cambridge University Press, Cambridge, England. Chapters 12 and 13 give the methods used and results obtained in deducing the nature of the interior of the earth from seismic data.

Clark, S. P. (1971). *Structure of the earth.* 132 pp. Prentice-Hall, Englewood Cliffs, N.J. A brief but comprehensive account utilizing geochemical, petrological, geothermal, and seismic data.

Clarke, F. W. and H. S. Washington (1924). The composition of the earth's crust. *U.S. Geol. Survey, Profess. Paper 127.* 117 pp. An exhaustive discussion and analysis of the data on the chemistry of the earth's crust.

Cloud, P. E. (1976). Beginnings of biospheric evolution and their biogeochemical consequences. *Paleobiology* **2**, 351–387.

Elsasser, W. M. (1963). Early history of the earth. In *Earth science and meteoritics,* pp. 1–30. John Wiley & Sons, New York. This paper introduces a novel but convincing hypothesis for the formation of the core by partial melting in a homogeneous earth.

Flanagan, F. J. (1974) Descriptions and analyses of seven new USGS rock standards. *U.S. Geol. Surv. Prof. Paper 840.* Contains an evaluation of the analytical data on the standard rocks G-1 and W-1; for earlier work see *U.S. Geol. Surv. Bull. 1113* (1960) and *980* (1951) and *Geochim. Cosmochim. Acta* **26**, 525–543(1962) and **29**, 1263–1283 (1965).

Hart, P. J. (ed.) (1969). *The earth's crust and upper mantle.* 735 pp. American Geophysical Union, Washington, D.C. A broad survey of modern developments, with emphasis on the interrelation of geochemical, geophysical, and geological observations.

MacDonald, G. J. F. (1964). Dependence of the surface heat flow on the radioactivity of the earth. *J. Geophys. Res.* **69**, 2933–2940. An examination of temperature variations within the earth as a result of radioactive disintegrations, and their geophysical and geochemical implications.

Ringwood, A. E. (1975). *Composition and petrology of the earth's mantle.* 618 pp. McGraw-Hill Book Co., New York. A comprehensive account, utilizing a broad array of evidence from several fields, including mineralogy, geochemistry, cosmochemistry, geophysics, and experimental data on silicate systems at high temperatures and pressures.

Robinson, E. C. (ed.) (1972) *The nature of the solid earth.* 677 pp. McGraw-Hill Book Co., New York. A symposium on the geochemistry and geophysics of the earth's interior.

Taylor, S. R. (1964). Abundance of chemical elements in the continental crust: a new table. *Geochim. Cosmochim. Acta* **28**, 1273–1285. A critical compilation and evaluation of the large amount of recent data on minor and trace elements in granites and basalts.

Wyllie, P. J. (1971). *The dynamic earth.* 416 pp. John Wiley & Sons, New York. A comprehensive account of the crust and mantle of the earth, with particular reference to recent discoveries.

4 SOME THERMODYNAMICS AND CRYSTAL CHEMISTRY

INTRODUCTION

In the introductory chapter geochemistry was described as dealing with the abundance, distribution, and migration of the chemical elements in the earth. The basic units of geochemical investigation are thus the elements, either in the form of atoms or more often as charged particles or ions. Atoms and ions have a certain energy content which changes when they undergo a physical or chemical transformation. In redistribution and recombination of the chemical elements in minerals and rocks the atoms or ions lose part of their energy and yield more stable systems. Every rock exemplifies the laws conditioning the stability of crystal lattices, laws that follow the general principles of the structure of matter and of thermodynamics. Geochemical concepts are completely meaningful when they indicate the relationship between atoms, ions, and crystal lattices and the factors determining their equilibrium conditions. Such relations properly belong to the field of chemistry, but in view of their significance in geochemistry a brief summary is given here.

FUNDAMENTAL THERMODYNAMIC EQUATIONS

The ideas of thermodynamics are most readily expressed in the form of equations. Only a brief summary of these equations is presented here; for further information reference should be made to a textbook of chemical thermodynamics or to the discussion of Wood and Fraser (1976) or Powell (1978) on the application of thermodynamics to geological processes. A useful source of thermodynamic properties of minerals and related substances is *U.S. Geological Survey Bulletin 1452* (Robie, Hemingway, and Fisher, 1978).

65

A thermodynamic system is characterized by certain fundamental properties, divisible into two types: (a) extensive or capacity properties, such as mass, volume, and entropy, which depend on the quantity of matter in the system; (b) intensive properties, such as temperature, pressure, and chemical potential, which are independent of the amount of matter in the system. The total energy of all kinds contained within a system is called its internal energy. (E). It depends only on the state of the system and cannot be determined in absolute values; it is the change in internal energy that the system undergoes in passing from one state to another which is significant.

The first law of thermodynamics states that energy can neither be created nor destroyed. If a system undergoes a change of state, and E_1 is the internal energy in the first state and E_2 the internal energy in the second state, then ΔE, the change in internal energy, is

$$\Delta E = E_2 - E_1$$

If in this change an amount of energy q is absorbed by the system in the form of heat and an amount of energy w leaves the system as mechanical work, then

$$\Delta E = q - w$$

For an infinitesimal change

$$dE = dq - dw$$

The mechanical work, dw, is usually measured by a change in volume dV acting against a hydrostatic pressure P, in which case

$$dw = P\,dV$$

so that

$$dE = dq - P\,dV$$

The second law of thermodynamics can be stated in the following form: "In any reversible process the change in entropy (dS) of a system is measured by the heat (dq) received by the system divided by the absolute temperature (T), that is, $dS = dq/T$; for any spontaneous irreversible process $dS > dq/T$." Thus, for a reversible process, the preceding equation can be restated in the form

$$dE = T\,dS - P\,dV$$

Because many processes take place at constant pressure with only heat energy and mechanical energy involved and because under these circumstances the heat energy absorbed by the system from its surroundings is equal to the increment in the ($E + PV$) function of the system, it has been found convenient to define a function H, called the enthalpy, such that

$$H = E + PV$$

Hence for any infinitesimal transformation

$$dH = dE + P\,dV + V\,dP$$

If the transformation occurs at constant pressure (i.e., $dP = 0$), then

$$dH = dq$$

that is, the change in enthalpy in any process at constant pressure is measured by the heat received and for this reason is often referred to as the heat of a reaction.

The Helmholtz free energy (A) and the Gibbs free energy (G) are defined by the following equations:

$$A = E - TS$$
$$G = E - TS + PV$$

The Gibbs free energy is especially significant in connection with processes that take place at constant temperature and pressure. Under these conditions

$$dG = dE - T\,dS + P\,dV$$

If the reaction is reversible

$$dE = T\,dS - P\,dV$$

and

$$dG = 0$$

Now a reversible reaction is synonymous with a state of equilibrium, and so we have as a criterion of equilibrium at constant pressure that the Gibbs free energy (hereafter referred to simply as the free energy) of the reactants must be equal to that of the products.

All geochemical processes may be regarded as striving toward equilibrium, which may be approached rather closely when composition, temperature, and pressure remain approximately constant for a long time. However, equilibrium if attained is seldom preserved, owing to changes in physical conditions. For example, if a thermally metamorphosed shale reached equilibrium under the conditions of metamorphism it is probably no longer in equilibrium at ordinary temperatures. However, the study of equilibria in laboratory experiments and by thermodynamic methods has thrown a flood of light on geochemical reactions, such as the origin of rocks and minerals, the processes of weathering and decomposition, and other kinds of transformations going on within the earth.

Any reaction, chemical or physical, may be characterized by an equation representing the transition from one state to another. If the reactants are in equilibrium with the products, ΔG for the reaction is zero. A large negative value of ΔG means that the reaction as written tends to proceed nearly to completion; a large positive value means that the reaction tends to proceed in the opposite direction. As a simple example, we may consider the solubility of a solid in a liquid; if the reaction is represented by the equation $S \rightarrow L$, then at saturation ΔG is zero, whereas a large negative value of ΔG means undersaturation, that is, a strong tendency to dissolve, and a positive value of ΔG signifies supersaturation.

Equations have been derived to express the effect of temperature and pressure on equilibrium. For temperature changes the relevant equation is

$$\frac{d(-\Delta G/T)}{dT} = \frac{\Delta H}{T^2}$$

This equation signifies that, if ΔH is positive, an increase in temperature makes ΔG more negative; i.e., if heat is absorbed in a reaction, an increase in temperature causes the reaction to go more nearly to completion. If ΔH is negative, on the other hand, increasing the temperature tends to inhibit the reaction. The effect of pressure is characterized by the equation $d\,\Delta G/dP = \Delta V$, where ΔV is the aggregate change in volume that occurs when the reaction proceeds to completion in the direction indicated. Thus, if ΔV is negative, an increase of pressure makes ΔG more negative; that is, the reaction as written proceeds more nearly to completion. In other words, high pressure favors the existence of materials of small volume, that is, high density. These two equations give quantitative expression to Le Châtelier's principle, which can be stated thus: if a system is in equilibrium, a change in any of the factors determining the conditions of equilibrium will cause the equilibrium to shift in such a way as to nullify the effect of this change. An interesting conclusion is that in general increasing temperature produces the same kind of effect as decreasing pressure. The volume of a substance becomes greater at higher temperatures and lower pressure. Solubility of solids in liquids increases as a rule with increasing temperature but is usually diminished by high pressure.

For a chemical reaction to proceed spontaneously, it is necessary that the total free energy of the products be less than that of the reactants, that is, $dG < 0$. The value of dG is a measure of the driving force or *affinity* of the reaction. To predict whether a certain reaction can occur it is therefore necessary to determine its free-energy change. The factors influencing the free energy change are (a) the composition of the reactants and of the products of reaction, (b) their state of aggregation, (c) the relative amounts present, and (d) the pressure and temperature. The experimental determination of the free-energy change of a chemical reaction is often extremely difficult. Nevertheless, if the discussion is confined to the standard states of the reactants and reaction products, and heat capacity data are available, the problem is often simplified, because the molal free energies of formation of many chemical compounds in their standard states are now known. Under these conditions the free-energy change is merely the sum of the free energies of the products of the reaction minus the sum of the free energies of the reacting substances. These free energies are obtained by multiplying the standard molal free energy of formation of each substance by the number of moles that enter into the reaction. Unfortunately, few data are yet available on the free energies of silicates, but this information is gradually being accumulated.

The major value of thermodynamics in geochemistry is that it provides a general approach to problems of stability, equilibrium, and chemical change. Even with qualitative data it enables predictions to be made regarding the probable course of all types of transformation. In any reaction for which the free energies of all possible phases are known under the specified conditions, thermodynamic equations permit the calculation of the relative amounts of reactants and products at equilibrium. If the amount of the products at equilibrium is found to be very small, then the reaction

is not favored under the specified conditions. If the amount of the products is large, the suggested reaction is one that may be expected to go under the specified conditions. In geochemical processes, many of which proceed under conditions that cannot be reproduced in the laboratory, thermodynamics provides us with the means for predicting the conditions under which certain reactions may occur, even though we cannot reproduce them experimentally. It is important to realize, however, that thermodynamics cannot predict the *rate* at which a reaction will proceed and does not tell us anything of the mechanism of the reaction.

One of the first practical applications of thermodynamics to the solution of a geochemical problem was a study of the stability of jadeite, $NaAlSi_2O_6$ (Kracek, Neuvonen, and Burley, 1951). Jadeite occurs in metamorphic rocks but at that time had never been made in the laboratory; its comparatively high density suggested that it might only be stable under high pressures. Kracek and his co-workers examined the thermodynamics of the following reactions by which jadeite might be formed:

$$NaAlSi_3O_8 = NaAlSi_2O_6 + SiO_2$$
$$NaAlSiO_4 + NaAlSi_3O_8 = 2NaAlSi_2O_6$$
$$NaAlSiO_4 + SiO_2 = NaAlSi_2O_6$$
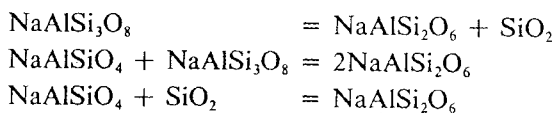
By measuring the heats of solution (in HF) of albite ($NaAlSi_3O_8$), nepheline ($NaAlSiO_4$), jadeite, and quartz, they were able to determine ΔH values for each of the above reactions. The specific heats of these substances had been measured and the standard entropies calculated, so it was then possible to evaluate ΔG from the equation $\Delta G = \Delta H - T\Delta S$. The figures obtained indicated that at 25° and 1 atm pressure the first reaction would tend to proceed from right to left (i.e., jadeite would not be formed), whereas the other two reactions would tend to proceed from left to right, with the formation of jadeite. These results show that jadeite is more stable at ordinary temperature and pressure than mixtures of albite and nepheline, or of nepheline and quartz, and that its stability is not conditioned by high pressure. The difficulties of making jadeite in the laboratory are therefore to be ascribed to kinetic factors involving activation energies and rates of reaction.

Even if all the relevant thermodynamic data are not available, Miyashiro (1960) has pointed out that it is frequently possible to derive useful information from heats of reaction alone, if the reactions involve only solids, a situation common under geological conditions. If all the reactants and products are in the solid state, the heat capacity changes are small and can be neglected. Under these circumstances the free energy $\Delta G_{T,P}$ of a reaction at temperature $T°K$ and pressure P atm is given by the following equation

$$\Delta G_{T,P} = \Delta H°_{298} - T\Delta S°_{298} + P\Delta V$$
$\Delta H°_{298}$ = heat of reaction at 298.16°K, i.e., 25°C, and 1 atm
$\Delta S°_{298}$ = entropy of reaction at 298.16°K and 1 atm

With all the phases in the solid state, $\Delta S°_{298}$ and ΔV are usually quite small in comparison with $\Delta H°_{298}$. Thus the heat of such a reaction at 25°C and 1 atm ($\Delta H°_{298}$) is nearly equal to the free energy of the same reaction at any temperature and pressure

$(\Delta G_{T,P})$. In other words, in reactions involving solid phases only, the free energy of reaction is usually nearly constant throughout a wide range of temperature and pressure. This is illustrated by the reaction:

$$Mg_2SiO_4 + SiO_2 = 2MgSiO_3$$

<center>forsterite   quartz   clinoenstatite</center>

$$\text{heat of reaction } \Delta H^\circ_{298} = -2300 \text{ cal/mole} \tag{1}$$

$$\text{entropy of reaction } \Delta S^\circ_{298} = -0.35 \text{ cal/deg mole}$$

$$\text{volume change } \Delta V = -4.0 \text{ cm}^3/\text{mole}$$

$$\text{free energy of reaction } \Delta G_{T,P} = -2300 + 0.35T - 0.097P$$

(the ΔV term must be divided by 41.3 to convert it to cal/mole).
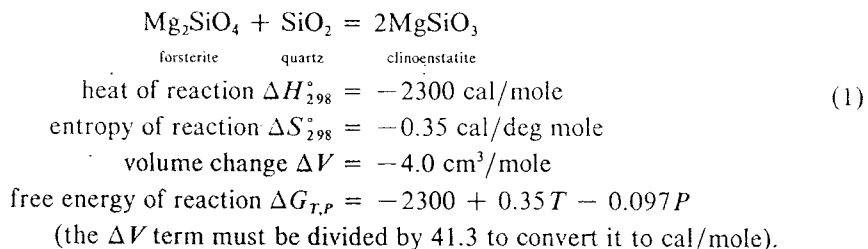
Clearly, the effects of temperature and pressure on the free energy of the reaction are very small. A change of temperature of 500° will change the free energy by 175 cal, and a change of pressure of 1000 atm will change it by less than 100 cal. There are, however, two points to be noted. First, even in the reactions involving solids only, the effect of differences in heat capacities cannot be neglected when temperature ranges are large. The upper limit for extrapolation from the standard state at 25° is probably about 500° in most cases. Second, under very high pressures such as in the mantle, the $P\Delta V$ term can no longer be overlooked. For example, for this reaction a change in pressure of 100 kb results in a change of the order of 10,000 cal in free energy.

Bearing in mind these limitations, let us consider further examples.

For the reaction

$$\tfrac{1}{2}NaAlSiO_4 + SiO_2 = \tfrac{1}{2}NaAlSi_3O_8 \tag{2}$$

<center>nepheline   quartz   albite</center>

$$\Delta G_{T,P} = -2500 - 0.05T + 0.016P$$

For the reaction

$$KAlSi_2O_6 + SiO_2 = KAlSi_3O_8 \tag{3}$$

<center>leucite   quartz   orthoclase</center>

$$\Delta G_{T,P} = -4800 + 1.46T - 0.053P$$
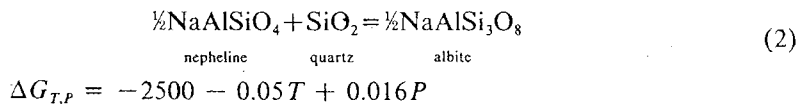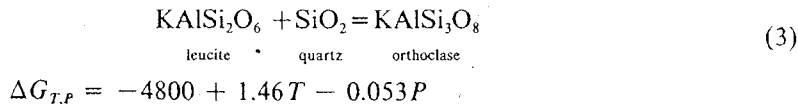
It is clear from these examples that $\Delta G_{T,P}$ cannot be greatly different from ΔH°_{298}.

We can use the available data to make deductions of petrogenetic significance. If the SiO_2 content of a rock consisting of clinoenstatite, albite, and orthoclase were gradually decreased, the preceding data indicate that the first reaction to take place would be the conversion of clinoenstatite into forsterite, since of the possible reactions this involves the smallest increase in free energy. This would be followed by the conversion of albite into nepheline, and finally by the conversion of orthoclase into leucite. The order of reaction thus predicted agrees with petrological experience. In rocks slightly undersaturated in SiO_2 olivine, $(Mg,Fe)_2SiO_4$, appears instead of pyroxene, $(Mg,Fe)SiO_3$. At a greater degree of undersaturation nepheline appears in place of albite, and, with a still greater degree of undersaturation, leucite is formed instead of orthoclase. This explains why nepheline is found in association with olivine

but not with $(Mg,Fe)SiO_3$ (hypersthene); the free energy relations favor the reaction of nepheline with hypersthene to give albite plus olivine.

However, the difference in the free energies of reaction (1) and reaction (2) above is not large, and high pressures, such as those at depth of 100–200 km, could be expected to reverse the order of the values of the free energies. Under these circumstances silica undersaturation might result in the conversion of albite into nepheline before the conversion of pyroxene into olivine. Considerations of this sort indicate that material of basaltic composition which crystallizes within the crust as a plagioclase-olivine rock might crystallize within the upper mantle as a nepheline-pyroxene rock. A unique advantage of thermodynamic data is that they facilitate the extrapolation of laboratory data to the more extreme conditions that can be expected in geological processes.

THE STATES OF MATTER

Geochemistry is to a large extent concerned with the transformation of matter from one state to another, as exemplified by the crystallization of magmas, the weathering of rocks, the deposition of salts from solution, and generally the formation of minerals over a wide range of temperatures, pressures, and chemical environments. These processes involve a change of state in all or part of the material. Three states of matter are recognized: solid, liquid, and gaseous. This division is a useful one, yet it should be realized that it is to some degree arbitrary. As we generally observe them, these states are sharply marked off from each other by distinctive properties, but under some conditions the boundaries lack definition, and the transitions solid ⇌ liquid ⇌ gas may be continuous rather than discontinuous.

In terms of the atomic theory, the state of matter ranges from complete atomic disorder in gases to complete order in crystals. However, complete order is an abstract concept and exists only in perfect crystals at absolute zero. At any temperature above absolute zero the kinetic energy of the atoms causes them to vibrate about their mean positions in the crystal lattice. If the kinetic energy of the atoms becomes sufficiently large, the crystal loses its rigidity; that is, it melts or decomposes. Usually, fusion takes place at a definite temperature, but theoretical considerations indicate that the melting point is a temperature range, possibly too small to measure but nevertheless finite in theory. Melting-point ranges have been observed experimentally. A special case of this phenomenon is illustrated by some complex organic substances that melt to give a liquid in which the molecules still maintain a one- or two-dimensional orientation, which is lost at still higher temperatures.

Liquids were formerly considered more akin to gases than to solids, and, indeed, beyond the critical point the distinction between liquid and gas ceases to have any validity. At temperatures and pressures well below the critical point, however, liquids may resemble solids. They diffract X-rays, and the diffraction effects show that a considerable degree of order exists in the arrangement of the atoms or molecules. A glass is simply a supercooled liquid held together by bonds extending throughout the structure, much as in the crystalline form of the same substance. The essential difference is that in the crystalline form the atoms are arranged in a symmetrical peri-

odic network, whereas in glasses the degree of orientation and periodicity is much lower.

All matter strives to reach equilibrium with its environment. To accomplish this the atoms try to arrange themselves in such a way that the free energy of the system is a minimum, and in the solid state this arrangement is usually an ordered crystal structure or structures. In geological terms, those minerals are formed that are most in harmony with the physical environment and the bulk composition of the system. Glasses are metastable phases, and even though they may persist for an almost indefinite period they always tend to change into crystalline forms. Thus glasses are uncommon in rocks, and their occurrence signifies unusual conditions of composition and formation.

THE CRYSTALLINE STATE

The most obvious characteristic of crystals that have grown freely is their external form. Morphological crystallography, the study of the geometrical relationships of the faces of crystals, has shown that every crystal can be classified into one of 32 classes based upon symmetry. Near the end of the eighteenth century, Haüy conceived that the geometrical complex of crystal faces characteristic of a homogeneous substance must be determined by its internal structure, the molecular or atomic arrangement. In 1912 the truth of Haüy's conception was demonstrated experimentally when the discovery of X-ray diffraction by crystals showed that in the crystalline state there is an orderly, systematic arrangement of atoms. The atomic arrangement largely determines the chemical and physical properties of a crystalline compound and is thus a fundamental feature.

PRINCIPLES OF CRYSTAL STRUCTURE

Since 1912 the crystal structures of many substances have been determined. Since minerals provide a ready source of well-crystallized substances, many of the early workers in this field naturally used them for crystal structure investigations. As a result, and fortunately for the progress of geochemistry, the structures of many minerals were worked out comparatively soon. A valuable summary of the data on the atomic structure of minerals has been provided by Bragg and Claringbull (1965). Goldschmidt and his associates also made important contributions in this field; it was one of Goldschmidt's major services to geochemistry that he early realized the significance of crystal structure in controlling the distribution of elements in the earth's crust.

The basic unit in all crystal structures is the atom (the term *atom* also includes ion in this discussion), which may, however, be associated with other atoms in a group behaving as a single unit in the structure. We may consider atoms as being made up of electric charges distributed through a small sphere that has an effective radius of the order of 1 Å (10^{-8} cm). The radius can be measured with considerable accuracy and depends not only on the nature of the element but also on its state of ionization and the manner in which it is linked to adjacent atoms. For example, the

radius of the sodium atom in metallic sodium is 1.86 Å, but the radius of the sodium ion in sodium salts is 1.02 Å.

The different kinds of interatomic linkage are classed into four bond types: the metallic bond responsible for the coherence of a metal; the ionic or polar bond, which is the linkage in salts such as sodium chloride; the covalent or coordinate link present in crystals such as the diamond; and the residual or van der Waals' bond, which is responsible for the coherence of the inert gases when condensed to solids at low temperatures. These four types of bonds all impart characteristic properties to the substances in which they occur and provide a convenient basis for the classification of crystal structures. More than one type of bond may occur in a single compound; such substances are termed *heterodesmic,* and those in which only one bond type is present *homodesmic.* In heterodesmic structures the physical properties, such as hardness, mechanical strength, and melting point, are in general determined by the weakest bonds, which are the first to suffer disruption under increasing mechanical or thermal strain.

It should always be realized that although these four types of bonding have well-defined properties the classification is arbitrary insofar as the bonding in many compounds may be more or less intermediate. The silicon-oxygen bonds in silica and the silicates are neither purely ionic nor purely covalent but are intermediate in nature. The structure assumed by any solid is such that the whole system of atomic nuclei and electrons tends to arrange itself in a form with minimum potential energy. The energy of a configuration can theoretically be calculated by applying the principles of quantum mechanics, and no distinction between the different types of bonds appears in the rigid mathematical expressions. Electronegativity, as discussed in Chapter 5, gives some indication of the ionic-covalent character of a chemical bond. Appendix I lists the estimated electronegativites for some of the elements. Cations with electronegativities furthest from the electronegativity of oxygen have the highest ionic bond character with oxygen.

Refinements of chemical bonding theory are necessary to explain some particular aspects of crystal chemistry. These include the assignment of electrons to particular atomic orbitals and the energy relationships between these orbitals and the formation of hybrid orbitals. Such considerations are useful, for example, in explaining the physical differences between graphite and diamond, both compounds of pure carbon. Carbon has six electrons in the unhybridized configuration.

$$1\,s^2\,2s^2\,2p_x^1\,2p_y^1$$
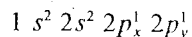
The $2p_z$ orbital is vacant. In graphite a $2\,sp^2$ hybridized orbital is formed from a 2-s, a $2p_z$, and a $2p_3$ orbital. This structure is planar with three lobes projecting at 120° from each other. The unused $2p_y$ orbital is perpendicular to $2sp^2$. The individual atoms are bonded to form planar sheets of carbon atoms. On the other hand, in diamond the four bonding electrons form a $2sp^3$ hybridized orbital that is made up of four equal orbitals in a tetrahedral orientation. Bonding of these leads to a three-dimensional structure of great strength. High pressures are necessary to make this structure the most stable configuration.

Crystal-field theory, related to the bonding of transition–metal elements with *d*

orbital electrons has proved to be useful in understanding and predicting the geo-chemistry of these important elements. Differences in geochemical behavior-between transition elements and other cations with similar charges and sizes may often be attributed to a crystal-field stabilization energy due to selective orientation of d orbital bonding electrons in the electrical fields formed by anions surrounding the cation site of interest in a crystal. For example, the divalent ions of Ti, V, Cr, Fe, Co, Ni, and Cu receive a stabilization energy in octahedral sites. Mn, Zn, and non-transition elements of similar size do not have this extra bond stabilization. Such considerations have been used to predict, for example, the stability and concentration of Fe, Co, and Ni in the lower mantle. The application of crystal-field theory to transition-element geochemistry is thoroughly covered by Burns (1970).

Nearly all the common minerals may be looked on as ionic structures, and we can consider them as compounds of oxygen anions with practically all the other elements (except the halogens) acting as cations. The oxygen ion is so large in comparison to most cations that a mineral structure is mainly a packing of oxygen ions with the cations in the interstices. The radii of common ions are given in the appendix and illustrated in Figure 4.1. Hydrogen is not included, because it has unique properties. The hydrogen ion or proton is so small that it can hardly be considered as having any spatial extension; instead it acts rather like a dimensionless center of positive charge. The radius of the OH^- ion is essentially the same as that of the O^{2-} ion; the hydrogen is embedded in the oxygen atom, and the OH group is effectively a sphere.

Because the radius of an ion depends on its atomic structure, it is related to the position of the element in the periodic table. The following rules are generally valid:

1. For elements in the same group of the periodic table, the ionic radii increase as the atomic numbers of the elements increase; e.g., Be^{2+} 0.27, Mg^{2+} 0.72, Ca^{2+} 1.00, Sr^{2+} 1.13, Ba^{2+} 1.36. This is, of course, to be expected, since for elements in the same group of the periodic table the number of electron orbits around the nucleus, and hence the effective radius, increases in going down the column.

2. For positive ions of the same electronic structure the radii decrease with increasing charge. For an example, we may take the elements in the second horizontal row in the periodic table, all of which have two electrons in the inner orbit and eight in the outer orbit:

$$\begin{array}{cccccc} Na^+ & Mg^{2+} & Al^{3+} & Si^{4+} & P^{5+} & S^{6+} \\ 1.02 & 0.72 & 0.53 & 0.40 & 0.17 & 0.12 \end{array}$$

Thus in going across a horizontal row in the periodic table the radii of the ions decrease. As electrons are lost the nucleus exerts a greater pull on those remaining, thus decreasing the effective radius of the ion.

3. For an element that can exist in several valence states, i.e., form ions of different charge, the higher the positive charge on the ion, the smaller the radius; for example, Mn^{2+} 0.82, Mn^{3+} 0.65, Mn^{4+} 0.54. The same reason given in the previous rule applies here also; the loss of an electron causes the remaining electrons to be more strongly attracted by the nucleus, thus effectively contracting the outer electron orbits and decreasing the ionic radius.
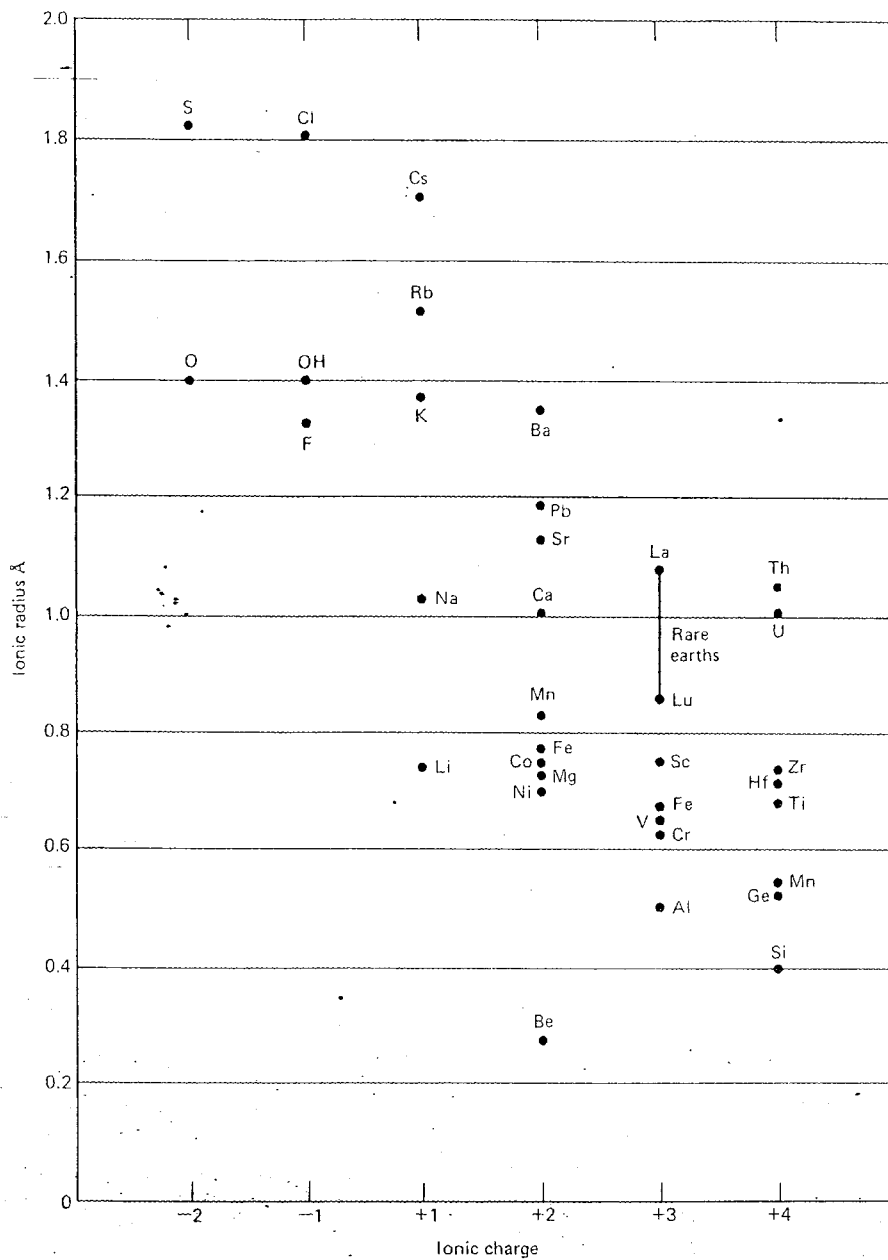
Figure 4.1 The relationship between ionic radius and ionic charge for some of the elements.

An apparent contradiction to the first rule is provided by the rare earth elements. The trivalent ions of these elements decrease in radius with increasing atomic number, from 1.14 for La^{3+} to 0.85 for Lu^{3+}. This remarkable feature, known as the lanthanide contraction, is the consequence of the building up of an inner electron shell, instead of the addition of a new shell; as a result the increasing nuclear charge produces an increased attraction on the outer electrons and an effective decrease in ionic radius. The lanthanide contraction also influences the geochemistry of the elements following lutetium; hafnium and tantalum have ionic radii almost identical with the elements above them in the periodic table—zirconium and niobium—and therefore show almost identical crystallochemical properties.

In an ionic structure each ion tends to surround itself with ions of opposite charge; the number that can be grouped around the central ion depends upon the radius ratio between the two. Figure 4.2 is a planar representation of the relationship. Assuming that ions act as rigid spheres of fixed radii, the stable arrangements of cations and anions for particular radius ratios can be calculated from purely geometric considerations (Table 4.1). Table 4.2 gives the radius ratio and predicted coordination number with respect to oxygen for the commoner cations, together with the coordination actually observed in minerals. The close correlation between observation and prediction confirms the assumption that ions act as spheres of definite radius; for cations larger than oxygen, however, the coordination number is less well defined, since the grouping is less regular.

Many cations occur exclusively in a particular coordination; others, for example,
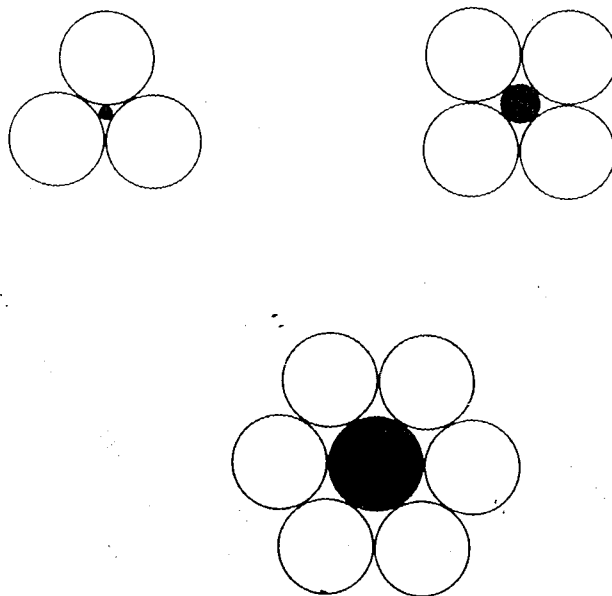
Figure 4.2 Planar representation of the relationship between radius ratio and coordination number.

Table 4.1 Relationship between Radius Ratio and Coordination Number for Ions as Rigid Spheres

Radius Ratio (R_{cation}/R_{anion})	Arrangement of Anions Around Cation	Coordination Number of Cation
0.15–0.22	Corners of an equilateral triangle	3
0.22–0.41	Corners of a tetrahedron	4
0.41–0.73	Corners of an octahedron	6
0.73–1	Corners of a cube	8
1	Closest packing	12

Table 4.2 The Relationship between Ionic Size and Coordination Number with Oxygen for the Commoner Cations

Ion	Radius (R)	R/R_0	Predicted Coordination Number	Observed Coordination Number
Cs^+	1.70	1.22	12	12
Rb^+	1.49	1.06	12	8–12
K^+	1.38	0.98	8	8–12
Ba^{2+}	1.36	0.97	8	8–12
Sr^{2+}	1.13	0.81	8	8
Na^+	1.02	0.73	6	6,8
Ca^{2+}	1.00	0.71	6	6,8
Mn^{2+}	0.82	0.59	6	6
Fe^{2+}	0.77	0.55	6	6
Li^+	0.74	0.53	6	6
Mg^{2+}	0.72	0.51	6	6
Ti^{4+}	0.67	0.48	6	6
Fe^{3+}	0.65	0.46	6	6
V^{3+}	0.64	0.46	6	6
Cr^{3+}	0.62	0.44	6	6
Al^{3+}	0.53	0.38	4	4,6
Si^{4+}	0.40	0.29	4	4

aluminum, which has a radius ratio lying near the theoretical boundary between two types of coordination, may occur in both. In such cases the effective radius depends to some degree on the coordination, increasing moderately with increasing coordination. The coordination is to some extent controlled by the temperature and pressure at which crystallization took place. High temperatures and low pressures favor low coordination, and low temperatures and high pressures favor higher coordination. High coordination is evidently more economical of space. Aluminum is a good example; in typically high-temperature minerals it tends to assume fourfold coordination and substitute for silicon, whereas in minerals formed at lower temperatures it tends to occur in sixfold coordination. The effect of high pressure on coordination has been dramatically illustrated by the synthesis of stishovite, a polymorph of SiO_2,

in which Si is in sixfold coordination instead of the fourfold coordination otherwise universal in silicate minerals.

The structure of ionic compounds is determined primarily by the demands of geometrical and electrical stability. The relative sizes of the ions and the mode of packing must result in the ions being more or less rigidly held in the structure, just as in a house built of blocks, where each block must support its neighbors. More than one structure may fulfill this requirement, in which case the most stable will be that for which the potential energy of the ions is lowest. The requirement of electrical stability means that the sum of positive and negative charges on the ions must balance. This is not achieved by pairing off individual cations and anions; the positive charge on a cation must be considered as divided equally between the surrounding anions, the number of which is determined by relative size of the ions and not by their valency. Pauling stated this in the following rule: "In a stable structure the total strength of the valency bonds which reach an anion from all the neighboring cations is equal to the charge on the anion." This expresses the tendency of any structure to assume a configuration of minimum potential energy, whereby the charges on the ions are as far as possible neutralized by their immediate neighbors.

This rule may appear more or less self-evident, but it is highly significant in the rigorous conditions it imposes on the geometrical configuration of a structure, especially of complex substances such as the silicates. It may be called the cardinal principle of mineral chemistry. It often explains the nonexistence of certain types of compounds, although their formulas would be quite possible according to the requirements of valency. For example, the feldspars have linked silicon-oxygen and aluminum-oxygen groups with Si and Al in fourfold coordination. An oxygen linked to Si and Al has a valency of only one-fourth left unsatisfied. This cannot be balanced by Mg or Fe in sixfold coordination, for their contribution would be at least one-third. It can, however, be balanced by large univalent and bivalent cations having coordination numbers of eight or more, since they can supply the necessary small fractions. Hence we find the feldspars are compounds of Ca, Na, and K and do not contain Mg or Fe.

The above principles are the basis for the crystal chemistry of minerals. They express the conditions for low potential energy and so for high stability. Only very stable compounds can occur as minerals; less stable compounds either do not form in nature or soon decompose. Artificial compounds have been made in which these general principles of crystal structure are not closely followed, but such substances are not found as minerals.

THE STRUCTURE OF SILICATES

In all silicate structures so far investigated (except those formed at extreme pressures) silicon lies between four oxygen atoms. This arrangement appears to be universal in these compounds, and the bonds between silicon and oxygen are so strong that the four oxygens are always found at the corners of a tetrahedron of nearly constant dimensions and regular shape, whatever the rest of the structure may be like. The different silicate types arise from the various ways in which these silicon-

oxygen tetrahedra are related to each other; they may exist as separate and distinct units, or they may be linked by sharing corners (i.e., oxygens). Silicate classification is based on the types of linkages, which are as follows:

1. Independent tetrahedral groups: in this type the silicon-oxygen tetrahedra are present as separate entities. The resultant composition is SiO_4, and a typical mineral is forsterite, Mg_2SiO_4. This division of the silicates is known as the *nesosilicates*.

2. Finite linked tetrahedral groups: in this type the silicon-oxygen tetrahedra are linked by the sharing of one oxygen between each two tetrahedra. If two tetrahedra are linked in this way, the resulting composition is Si_2O_7; a typical mineral is akermanite, $Ca_2MgSi_2O_7$, and such substances are classed as *sorosilicates*. If more than two tetrahedra are so linked, closed units of a ring-like structure are formed, giving compositions Si_nO_{3n}. Rings containing up to six silicons are known. Typical examples are benitoite, $BaTiSi_3O_9$, with three linked tetrahedra, and beryl, $Be_3Al_2Si_6O_{18}$, with six. This division of the silicates is known as the *cyclosilicates*.

3. Chain structures: tetrahedra joined together to produce chains of indefinite extent. There are two principal modifications of this structure yielding somewhat different compositions: (a) single chains, in which $Si:O$ is $1:3$, characterized by the pyroxenes, and (b) double chains, in which alternate tetrahedra in two parallel single chains are cross-linked and the $Si:O$ ratio is $4:11$, characterized by the amphiboles. These chains are indefinite in extent, are elongated in the c-direction of the crystal, and are bonded to each other by the metallic elements. This division of the silicates is known as the *inosilicates*.

4. Sheet structures: three oxygens of each tetrahedron are shared with adjacent tetrahedra to form extended flat sheets. This is the double-chain inosilicate structure extended indefinitely in two directions instead of just one. This linkage gives a ratio of $Si:O$ of $2:5$ and is the fundamental unit in all mica and clay structures. The sheets form a hexagonal planar network responsible for the principal characteristics of minerals of this type—their pronounced pseudohexagonal habit and perfect basal cleavage parallel to the plane of the sheet. This division of the silicates is known as the *phyllosilicates*.

5. Three-dimensional networks: every SiO_4 tetrahedron shares all its corners with other tetrahedra, giving a three-dimensional network in which the $Si:O$ ratio is $1:2$. The various forms of silica—quartz, tridymite, cristobalite—have this arrangement. The quadrivalent silicon balanced by two bivalent oxygen atoms. In silicates of this type the silicon is partly replaced by aluminum so that the composition is $(Si,Al)O_2$. The substitution of Al^{3+} for Si^{4+} requires additional positive ions in order to restore electrical neutrality. The feldspars and zeolites are examples of this division of the silicates, which is known as the *tektosilicates*.

All the silicate minerals can be placed in one of the types listed in Table 4.3 (a few have more than one type of linkage in the structure).

The other constituents of a silicate structure, such as additional oxygen atoms, hydroxyl groups, water molecules, and cations, are arranged with the silicate groups

Table 4.3   The Structural Classification of the Silicates

Classification	Structural Arrangment	Silicon: Oxygen Ratio	Examples
Nesosilicates	Independent tetrahedra	1:4	Forsterite, Mg_2SiO_4
Sorosilicates	Two tetrahedra sharing one oxygen	2:7	Akermanite, $Ca_2MgSi_2O_7$
Cyclosilicates	Closed rings of tetrahedra each sharing two oxygens	1:3	Benitoite, $BaTiSi_3O_9$ Beryl, $Al_2Be_3Si_6O_{18}$
Inosilicates	Continuous single chains of tetrahedra each sharing two oxygens	1:3	Pyroxenes, e.g., enstatite, $MgSiO_3$
	Continuous double chains of tetrahedra sharing alternately two and three oxygens	4:11	Amphiboles, e.g., anthophyllite, $Mg_7(Si_4O_{11})_2(OH)_2$
Phyllosilicates	Continuous sheets of tetrahedra each sharing three oxygens	2:5	Talc, $Mg_3Si_4O_{10}(OH)_2$ Phlogopite, $KMg_3(AlSi_3O_{10})(OH)_2$
Tektosilicates	Continuous framework of tetrahedra each sharing all four oxygens	1:2	Quartz, SiO_2 Nepheline, $NaAlSiO_4$

in such a way as to produce a mechanically stable and electrically neutral structure. Aluminum, after silicon the most abundant cation in the earth's crust, plays a unique role. As discussed earlier, it is stable both in fourfold and in sixfold coordination. It can replace silicon in the SiO_4 groups and also the common six-coordination cations—Mg^{2+}, Fe^{2+}, Fe^{3+}, etc.

The valence charge on the silicate unit, which determines the number and charge of the other ions that may enter the structure, can easily be calculated if it is remembered that each silicon has a positive charge of four and each oxygen a negative charge of two. Thus, the charge on a single SiO_4 unit is $[4 + 4(-2)] = -4$; on an Si_2O_7 unit, -6; on an SiO_3 unit, -2; on an Si_4O_{11} unit, -6; on an Si_2O_5 unit, -2; and on SiO_2, 0.

THE LATTICE ENERGY OF CRYSTALS

It has been pointed out that the structure of a crystal is determined by the tendency of the constituent atoms to take up positions whereby their total potential energy is reduced to a minimum. This tendency may be expressed in terms of lattice energies. The lattice energy of an ionic crystal, generally represented by U, is defined as the energy absorbed when a mole of the crystal is dispersed into infinitely separated ions. The lattice energy depends on the balancing of (a) the electrostatic forces between ions of opposite charge, which give a resultant attraction falling off with the square of the distance, and (b) the internuclear repulsive forces, which fall off very rapidly with distance. The attractive and repulsive forces result in an equilibrium position of minimum potential energy, which summed over all the ions is numerically equal to

the lattice energy of the crystal. The greater the lattice energy the greater the energy required to break up the crystal into its constituent ions.

For binary compounds lattice energies may be directly calculated from the properties of the ions by means of the following equation, originally derived by Born and Landé:

$$U = \frac{NAz_cz_a}{r}\left(1 - \frac{1}{n}\right)$$

where U = lattice energy.

   N = Avogadro number.

   A = Madelung constant, characteristic of the type of crystal structure.

z_c, z_a = charge on cation and anion, respectively.

   r = shortest anion-cation distance.

   n = a factor allowing for the internuclear repulsion (n is usually about 10).

From the form of the equation it is clear that U approaches zero as r approaches infinity.

Although the Born-Landé equation applies only to binary compounds, it does enable qualitative statements regarding the lattice energy of more complex substances. For a particular type lattice energies are greater the higher the charge on the ions, the smaller the ions, and the closer the packing. For an example of the first effect, we may cite two substances with the same crystal structure, NaCl ($U = 183$ Cal) and MgO ($U = 939$ Cal).

The energy U of the Born-Landé equation is equal to the amount of work per mole that must be expended to disperse the crystalline substance into an assemblage of widely separated ions. As such, it cannot be equated with any directly measurable quantity and is not to be identified either with the heat of sublimation, which is the energy necessary to disperse the substance into a molecular gas, or with the heat of solution, which also includes the heat of hydration of the ions, or with the heat of formation, which is the heat evolved by the formation of the substance from its elements. Born and Haber devised a thermochemical cycle by means of which the lattice energy can be related to measurable thermal data. This cycle is as follows, using NaCl as an example (square brackets indicate crystalline substances and parentheses indicate substances in the gaseous state):

$$[NaCl] \xrightarrow{U} (Na^+) + (Cl^-)$$
$$\uparrow -Q \qquad\qquad\qquad \downarrow -I+E$$
$$[Na] + (\tfrac{1}{2}Cl_2) \xleftarrow{-S-D} (Na) + (Cl)$$
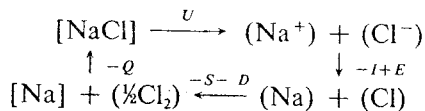
The diagram represents the following cycle:

1. One mole of the crystalline substance is dispersed into ions in the gaseous state.
2. The ions are converted into neutral atoms.
3. The neutral atoms, now in the form of monatomic gases, are converted into the standard states for the elements at 25° and 1 atm pressure.
4. The elements are then allowed to combine chemically to reform the crystalline substance.

The symbols are defined by the following thermochemical equations:

$$[NaCl] = (Na^+) + (Cl^-)- -\Delta H = U, \text{ lattice energy}$$
$$(Na^+) + e = (Na) \qquad \Delta H = I, \text{ ionization energy}$$
$$(Cl^-) = (Cl) + e \qquad \Delta H = E, \text{ electron affinity}$$
$$(Na) = [Na] \qquad \Delta H = S, \text{ heat of sublimation}$$
$$(Cl) = (\tfrac{1}{2}Cl_2) \qquad \Delta H = D, \text{ heat of dissociation}$$
$$[Na] + (\tfrac{1}{2}Cl_2) = [NaCl] \qquad \Delta H = Q, \text{ heat of formation}$$
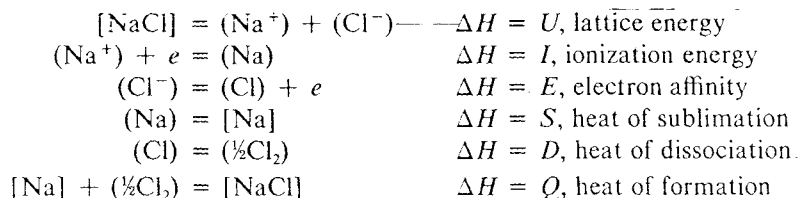
Since the final state of the system is the same as its original state, the net change in heat content is zero; hence

$$U = Q + S + I + D - E$$

The cycle given by this equation is to be regarded as isothermal at 25°.

The Born-Haber cycle thus provides us with a means of determining lattice energies for complex compounds from other thermodynamic quantities. Unfortunately, for the common silicate minerals these thermodynamic quantities are imperfectly known. The whole question of considering geochemical processes in terms of lattice energies has been a subject of particular interest to Fersman and other Russian geochemists. In an attempt to get a simple method of determining lattice energies Fersman introduced the *EK* concept, which is an empirical constant for each element, representing the contribution of that element to the lattice energies of its compounds. Thus for NaCl

$$U_{NaCl} = K(EK_{Na} + EK_{Cl}) \ (K \text{ is an independent constant})$$

From known lattice energies Fersman was able to assign *EK* values to most of the elements. However, the application of these values to the calculation of lattice energies of silicate minerals and to the interpretation of geochemical processes has so far given ambiguous results. The differences in calculated lattice energies of corresponding amphiboles and pyroxenes, for example, are small, smaller than the probable error. One serious criticism of Fersman's approach is that it allows only for the energy associated with a specific number of ions of different elements but fails to consider the energy associated with the crystal lattice as such. For example, the preceding equation will clearly give the same lattice energies for all polymorphs of the same substance.

ISOMORPHISM

The term *isomorphism* is applied to the phenomenon of substances with analogous formulas having closely related crystal structures. The term was introduced by Mitscherlich in 1819, who prepared crystals of KH_2PO_4, KH_2AsO_4, $(NH_4)H_2PO_4$, and $(NH_4)H_2AsO_4$ and found that they showed the same forms and the interfacial angles between corresponding faces were very similar. By Mitscherlich's original definition, substances with analogous formulas and similar crystallography said to be isomorphous. X-ray studies have shown that similar crystallography is a reflection of similar internal structure, hence the rewording of his original definition; sometimes the term *isostructural* or *isotypic* is used for the phenomenon.
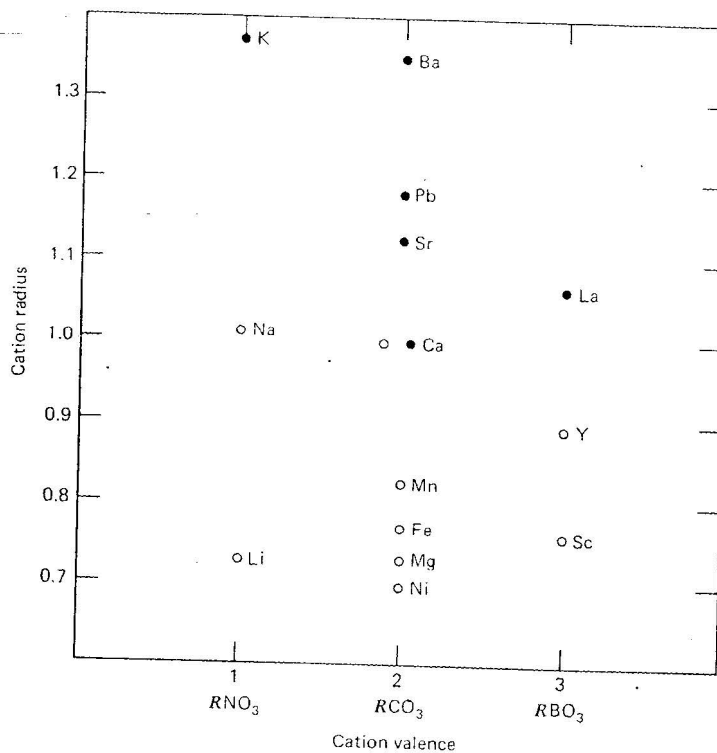
Figure 4.3 Effect of cation radius in determining crystal structure;
○ = trigonal; ● = orthorhombic. (After Fleischer, *J. Chem. Educ.* **31**, 450, 1954)

Isomorphism is widespread among minerals and is one of the bases of their classification. Many isomorphous groups are recognized, for example, the spinel group, the garnet group, and the amphibole group. The basis of the phenomenon is that anions and cations of the same relative size (i.e., showing the same coordination) and in the same numbers tend to crystallize in the same structure type. This is well exemplified by some of the carbonate minerals (Figure 4.3). The anhydrous carbonates of the bivalent elements form two isomorphous groups, one orthorhombic and one trigonal It can be seen that the nature of the structure is determined by the size of the bivalent cation; those minerals with cations larger than calcium crystallize in an orthorhombic structure, those with cations smaller than calcium crystallize in a trigonal structure. Calcium carbonate itself can crystallize in either structure, the phenomenon known as *polymorphism.*

Other substances with analogous formulas are isomorphous with these carbonates. Thus soda niter, $NaNO_3$, is isomorphous with calcite, whereas niter, KNO_3, is isomorphous with aragonite, reflecting the similar size of nitrate and carbonate groups and the larger size of potassium ions as compared to sodium ions. The borates of trivalent elements show similar relationships. Until the development of X-ray techniques for the determination of crystal structures, it was somewhat of an enigma that

substances as different chemically as calcite and soda niter could show complete similarity in crystal form. Other isomorphous pairs at first sight do not even have analogous formulas. Thus the rare mineral berlinite ($AlPO_4$) is isomorphous with quartz; the true analogy is seen when the formula of quartz is written $SiSiO_4$. Both Al and P are similar in ionic size to Si and can exist in a crystal structure in four-coordination with oxygen; as a result $AlPO_4$ can crystallize with the same structure as quartz. Similarly, tantalite, $FeTa_2O_6$, is isomorphous with brookite, $TiO_2(TiTi_2O_6)$; the metallic ions are similar in size and all show sixfold coordination with oxygen.

The important factor in isomorphism is the similarity in size relations of the different ions rather than any chemical similarity. This explains many apparently unusual examples of isomorphism and its absence between many chemically similar compounds. Thus corresponding calcium and magnesium compounds are seldom isomorphous, although these elements are similar in chemical behavior; when it is noted that the radius of Ca^{2+} is 1.00 Å and that of Mg^{2+} is 0.72 Å it seems natural that the substitution of one for the other without producing a change in structure is improbable.

ATOMIC SUBSTITUTION

After the development of reliable methods for the analysis of minerals it was observed that many species are variable in composition. Substitution of one element by another is the rule rather than the exception. When this phenomenon was first observed it was described in terms of the concept of *solid solution* or *mixed crystals*, which implied the presence, in a single homogeneous crystal, of molecules of two or more substances. For example, common olivine may be described as a solid solution of Mg_2SiO_4(Fo) and Fe_2SiO_4(Fa), and the precise composition of any sample of olivine may be stated in terms of these *end-members*, such as $Fo_{85}Fa_{15}$, that is, $(Mg_{0.85}Fe_{0.15})_2SiO_4$. This concept and terminology remain in general use, but the light thrown upon the structure of crystals by X-ray investigation has resulted in a revised interpretation. In an ionic structure there are no molecules, the structure being an infinitely extended three-dimensional network. Any ion in the structure may be replaced by another ion of similar radius without causing any serious distortion of the structure, just as a bricklayer, running short of red bricks, may incorporate yellow bricks of the same size here and there in his wall. Since minerals usually crystallize from solutions containing many ions other than those essential to the mineral, they often incorporate some foreign ions in the structure.

A solid solution (or mixed crystal) can be simply defined as a homogeneous crystalline solid of variable composition. It was early found that many isomorphous substances have the property of forming solid solutions. There has been a tendency to equate isomorphism and solid solution, in spite of marked inconsistencies. For example, many isomorphous substances show little or no solid solution, among them, calcite and smithsonite; and extensive solid solution may occur between components that are not isomorphous, for example, the presence of considerable amounts of iron in sphalerite, although FeS and ZnS have quite different crystal structures. On this account it must be emphasized that isomorphism is neither necessary to nor sufficient

for solid-solution formation. Isomorphism and solid solution are distinct concepts and should not be confused.

In atomic substitution it is the size of the atoms or ions that is the governing factor, and it is not essential that the substituting ions have the same charge or valency, provided that electrical neutrality is maintained by concomitant substitution elsewhere in the structure. Thus, in passing from albite ($NaAlSi_3O_8$) to anorthite ($CaAl_2Si_2O_8$), Ca^{2+} substitutes for Na^+ and electrical neutrality is maintained by the coupled substitution of Al^{3+} for Si^{4+}; similarly, in diopside ($CaMgSi_2O_6$), Mg^{2+}—Si^{4+} may be replaced in part by Al^{3+}—Al^{3+}. Such coupled substitutions are especially common in silicate minerals and made the interpretation of their composition exceedingly difficult before this phenomenon was recognized and understood.

As a general rule, little or no atomic substitution takes place when the difference in charge on the ions is greater than one, even when size is appropriate (e.g., Zr^{4+} does not substitute for Mn^{2+}, nor Y^{3+} replace Na^+); this may be in part because of the difficulty in balancing the charge requirements by other substitutions.

The extent to which atomic substitution takes places is determined by the nature of the structure, the closeness of correspondence of the ionic radii, and the temperature of formation of the substance. The nature of the structure evidently has considerable influence on the degree of atomic substitution; some structures, such as those of spinel and apatite, are well known for extensive atomic substitution, whereas others, such as quartz, show very little. In part this phenomenon is due to the lack of foreign ions of suitable size and charge. Ionic size has, of course, a fundamental influence on the degree of substitution, since the substituting ion must be able to occupy the lattice position without causing distortion of the structure. From a study of many mixed crystals it has been found that, provided the radii of substituting and substituted ions do not differ by more than 15%, a wide range of substitution may be expected at room temperature. Higher temperatures permit a somewhat greater tolerance; in this respect solid solutions are analogous to solutions of salts in water, solubility increasing with temperature.

This property of increased atomic substitution at higher temperatures provides a means of determining the temperature of mineral formation (*geological thermometry*). If for a specific mineral the degree of atomic substitution has been determined for different temperatures, the composition of the naturally occurring mineral may indicate the temperature of formation. Thus the amount of titanium in solid solution in magnetite as a function of temperature is known from laboratory investigations. Magnetite is a common ore mineral. Provided the ore-forming solutions contained sufficient titanium to saturate the magnetite, the titanium content of the mineral will indicate the temperature conditions during ore deposition.

The consequence of atomic substitution is that most minerals contain not only the elements characteristic of the particular species, but also other elements able to fit into the crystal lattice. For instance, dolomite is theoretically a simple carbonate of magnesium and calcium, but dolomites are found whose analyses show a considerable content of iron and manganese. Traditionally, these were described as solid solutions of the carbonates of all these elements, but it is more illuminating as well as more correct to consider them as products of the substitution of iron and man-

ganese for magnesium. Nevertheless, we continue to use the traditional terms solid solution, mixed-crystals, and solid solution series, since the terminology of atomic substitution has not yet provided expressions to take their place. The useful term *diadochy* has been introduced to describe the ability of different elements to occupy the same lattice position in a crystal; thus Mg, Fe, and Mn are diadochic in the structure of dolomite. The concept of diadochy, if used rigidly, always applies to a particular structure; two elements may be diadochic in one mineral and not in another.

Crystal structure investigations have also revealed two other types of solid solution besides that due to atomic substitution. One is known as *interstitial solid solution,* whereby atoms or ions do not replace atoms or ions in the structure but fit into interstices in the lattice. This type is very common in metals, which take up hydrogen, carbon, boron, and nitrogen, all small atoms, in interstitial solid solution. If a substance has an open structure, interstitial solid solution may take place even with atoms or ions of a considerable size. Thus cristobalite, the high-temperature form of SiO_2, has been found with a considerable content of sodium and aluminum; the Al^{3+} replaces Si^{4+}, and the Na^+ needed to maintain electrical neutrality occupies large openings in the cristobalite lattice. The other type of solid solution is that associated with *defect lattices,* in which some of the atoms are missing, leaving vacant lattice positions. It has been called *omission solid solution.* A good example is the mineral pyrrhotite, in which analyses always show more sulfur than corresponds to the formula FeS. This was for a long time described as solid solution of sulfur in FeS. Actually, the excess of sulfur shown by analyses is due to the absence of some iron atoms from their places in the lattice; there is a deficiency of Fe, not an excess of S. Just as in building a wall, where a brick may be omitted here and there without affecting the stability of the structure, so it is possible to omit some of the Fe atoms in FeS without the lattice collapsing. More and more defect structures are being recognized among minerals, thereby explaining otherwise puzzling deviations of chemical compositions from those predicted by the law of constant proportions.

POLYMORPHISM

An element or compound that can exist in more than one crystal form is said to be polymorphous. Each form has different physical properties and a distinct crystal structure; that is, the atoms or ions are arranged differently in different polymorphs of the same substance. Polymorphism is an expression of the fact that crystal structure is not exclusively determined by chemical composition, that there is often more than one structure into which the same atoms or ions in the same proportions may be built up. Different polymorphs of the same substances are formed under different conditions of pressure, temperature, and chemical environment; hence the presence of one polymorph in a rock will often tell something about the conditions under which that rock was formed. For example, marcasite can be formed only from acid solutions at temperatures below 450°, and the presence of marcasite in a deposit thus puts some limits on the conditions of origin.

Two types of polymorphism are recognized, according to whether the change from

one polymorph to another is reversible and takes places at a definite temperature and pressure. The first type is know as *enantiotropy* and is exemplified by the relationship between quartz and tridymite (quartz $\overset{867°}{\underset{1 \text{ atm}}{\rightleftharpoons}}$ tridymite). The second type is known as *monotropy;* an example is the marcasite-pyrite relationship, in which marcasite may invert to pyrite but pyrite never changes to marcasite. With monotropic polymorphs one form is always inherently unstable and the other inherently stable; the unstable form always tends to change into the stable form, but the stable form cannot be changed into the unstable form without first completely destroying its structure by melting, vaporization, or solution.

This distinction between enantiotropic and monotropic polymorphs is useful, but the recognition of monotropic polymorphs is usually based on experimental evidence, and investigation over wide ranges of temperature and pressure or determination of energy relationships of the different polymorphs sometimes indicates that supposedly monotropic polymorphs actually have an enantiotropic relationship under conditions far removed from those usually attainable. The diamond-graphite relationship is particularly interesting in this respect, both from the geological significance of the occurrence of these two polymorphs and from the practical aspect of developing ways to make diamond synthetically. For a long time it was unknown whether diamond and graphite were enantiotropic or monotropic polymorphs; the latter conclusion was favored because under laboratory conditions the transition was always diamond → graphite, never the reverse. However, it has been established that the relationship is enantiotropic, and the actual conditions of the diamond ⇌ graphite equilibrium have been worked out (Figure 4.4). This figure shows that the practical problem of making diamond synthetically lies in maintaining pressures within the stability field of diamond at temperatures for which the reaction velocity for its formation is appreciable; this has been achieved by the development of special equipment capable of withstanding great pressures at high temperatures. It also indicates that the natural occurrence of diamond in igneous rocks implies an origin at considerable depths in
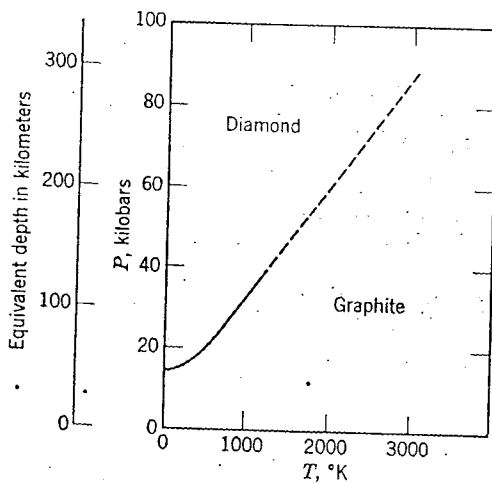
Figure 4.4 Graphite-diamond equilibrium curve, calculated to 1200°K, extrapolated beyond. (*Nature* **176**, 835, 1955)

the earth, where the combination of temperature and pressure is within the diamond stability field. Diamond is actually unstable under the physical conditions in which it is found (and worn); that it does not change spontaneously into graphite is due soley to the infinitesimal rate of a reaction which the energy relations nevertheless favor.

The rate of change from one polymorph to another may be very slow or very rapid. Sometimes the change does not involve the breaking of bonds between neighboring atoms or ions, but simply their bending, for example, low-quartz $\rightleftharpoons$ high-quartz, low-leucite $\rightleftharpoons$ high-leucite. Such transformations are almost instantaneous at the transition temperature, and the high-temperature form cannot be preserved at lower temperatures (however, original crystallization as the high-temperature form can often be recognized from the nature of the crystals or from the twinning that so often results from inversions of this type). High-low polymorphs are also characterized by the fact that the high-temperature form is always more symmetrical than the corresponding low-temperature form. Transformations other than the high-low type require the breaking of bonds in the structure and the rearrangement of atomic or ionic linkages. They are often sluggish and may require the presence of a solvent in order to obtain an appreciable rate of change. These changes have been termed reconstructive transformations and are exemplified by the quartz $\rightleftharpoons$ tridymite $\rightleftharpoons$ cristobalite inversions.

The high-temperature polymorph of a substance is generally more open-packed than a low-temperature form. The open character of the structure is dynamically maintained at high temperatures by thermal agitations. It may also be statically maintained by the incorporation of foreign ions in the interstices of the lattice. These foreign ions will buttress the structure and prevent its transformation to a different polymorph when the temperature is lowered. Their complete removal is usually necessary to permit inversion to the closepacked form stable at low temperatures. Thus impure high-temperature polymorphs may be formed and may survive indefinitely far below the normal stability range of pure compounds. This situation is likely to arise in nature. This phenomenon is probably responsible for the formation and survival of cristobalite and tridymite under conditions in which the stable form of SiO_2 is quartz. As mentioned previously, cristobalite has been found with a considerable amount of sodium in interstitial solid solution, and the sodium atoms presumably stabilize the open structure of this mineral. The occurrence of a high-temperature polymorph at ordinary temperatures is not necessarily to be interpreted as indicating metastability; the polymorph may be simply a stable impure form.

Transformations between polymorphs show a close analogy to the changes between the liquid and solid states, being amenable to treatment by the same thermodynamic principles. Under a given set of conditions, each of several polymorphs of a substance is characterized by its free energy; all possible polymorphic forms tend to transform into the one with the minimum free energy, and that form is the stable one under those conditions. The free energy G is given by the equation $G = E - TS + PV$. For changes not involving a gas phase the PV term is small and can be omitted. When T is zero, $G = E$, that is, the free energy of a substance is equal to
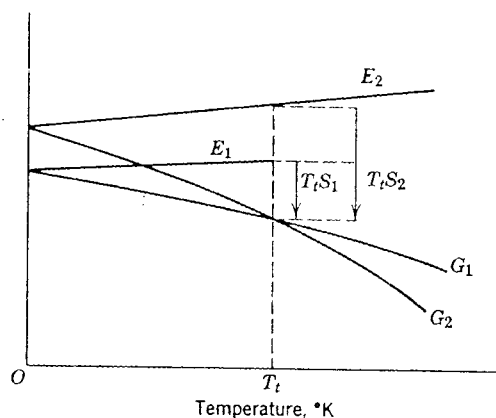
Figure 4.5 Energy relations between polymorphs 1 and 2 with transition temperature T_t. E = internal energy; G = free energy; S = entropy. (After Buerger, *Am. Mineralogist* **33**, 103, 1948)

its internal energy. Hence at absolute zero the polymorph with the lowest internal energy will be the stable form. At temperatures other than zero the entropy term is significant, and the relative magnitudes of S for the different polymorphs may determine which form has the lowest free energy. At a transition point between two forms, the free energies of both are equal (Figure 4.5). Because the entropy of the high-temperature form is greater than that of the low-temperature form, it follows that the internal energy of the former must also be greater than that of the latter. The entropy involves the volume over which the atoms may be disordered; hence there is a tendency for the forms of higher entropy to have greater open space available for thermal motion. Although this does not necessarily involve openness of the whole structure, it often does; high-temperature forms therefore tend to be less dense than low-temperature ones. Temperature and pressure tend to impose opposite conditions: high temperature promotes open structures, high pressures compact structures.

An interesting transformation that may be considered as a variety of polymorphism is the order-disorder type. It has been most studied in alloys, as it has important effects on their physical properties, but it is probably more common in minerals than is generally realized. A simple example is an alloy of 50% Cu, 50% Zn. Two distinct phases of this alloy exist. In the disordered form the copper atoms and the zinc atoms are randomly distributed over the lattice positions, whereas in the ordered form each element occupies a specific set of positions. The structures of the two forms are related, but the ordered one has lower symmetry than the disordered one. There is no definite transition point between the two forms; perfect order is achieved only at absolute zero, and with increasing temperature the degree of order gradually decreases to complete disorder above a certain temperature characteristic of the

structure and the composition of the crystal. For example the relationship between microcline and sanidine is an order-disorder transformation, the one aluminum and three silicon atoms in $KAlSi_3O_8$ being in disorder in sanidine but ordered in microcline. This accounts for the monoclinic-triclinic feature of the polymorphism, and the typical twinning of microcline is characteristic of the twinning often observed in ordered forms.

SELECTED REFERENCES

Bloss, F. D. (1971). *Crystallography and crystal chemistry.* 545 pp. Holt, Rinehart and Winston, New York. A comprehensive account of the principles of crystal chemistry, with special reference to their applications to mineralogy.

Bragg, W. L. and G. F. Claringbull (1965). *The crystal structures of minerals.* 409 pp. Bell and Sons, London. Chapter 2 is an excellent presentation of the factors governing the structures of minerals, and the following chapters provide a succinct account of mineral structures determined up to 1963.

Broecker, W. S. and V. M. Oversby (1971). *Chemical equilibria in the earth.* 318 pp. McGraw-Hill Book Co., New York. A comprehensive account of the application of thermodynamics and kinetics to chemical systems of geological importance.

Buerger, M. J. (1951). Crystallographic aspects of phase transformations. *Phase transformations in solids,* pp. 183–211. John Wiley & Sons, New York. An account of the thermodynamic background to polymorphism.

Burns, R. G. (1970). *Mineralogical applications of crystal field theory,* 224 pp. Cambridge University Press. Cambridge, England. A summary of crystal field theory and its application to the geochemistry of the transition elements.

Fyfe, W. S. (1964). *Geochemistry of solids.* 199 pp. McGraw-Hill Book Co., New York. A clear and concise account of the general principles of chemical bonding and related topics.

Kern, R. and A. Weisbrod (1967). *Thermodynamics for geologists.* 304 pp. Freeman, Cooper & Co., San Francisco. An account of basic thermodynamics of particular significance to geologists, and their application to a variety of geological processes and materials.

Kracek, F. C., K. J. Neuvonen, and G. Burley (1951). Thermochemistry of mineral substances, I: A thermodynamic study of the stability of jadeite. *J. Wash. Acad. Sci.* **41**, 373–383. A pioneer application of thermodynamical principles and measurements to the study of mineral formation.

Miyashiro, A. (1960). Thermodynamics of reactions of rock-forming minerals with silica. *Japanese J. Geol. Geog.* **31**, 71–78. A discussion of the free energies of solid-state reactions and their petrogenetic significance.

Pauling, L. (1960). *The nature of the chemical bond* (third ed.). 644 pp. Cornell University Press, Ithaca, N.Y. An exhaustive account of the types of bonding in elements and compounds and their significance for crystal structure and chemical properties.

Powell, R. (1978). *Equilibrium thermodynamics in petrology,* 283 pp. Harper & Row, New York. Thermodynamics of rock and fluid systems.

Robie, R. A., B. S. Hemingway, and J. R. Fisher (1978). Thermodynamic properties of minerals and related substances at 298.15 K and 1 bar (10^5 Pascals) pressure and at higher temperatures. 445 pp. *U.S. Geological Survey Bulletin 1452.*

Wood, B. J. and D. G. Fraser (1976). *Elementary thermodynamics for geologists,* 301 pp. Oxford University Press, Oxford. The application of thermodynamics to geochemical and petrological problems.

5 MAGMATISM AND IGNEOUS ROCKS

WHAT IS MAGMA?

Magma is the parent material of igneous rocks. Lava is magma poured out through volcanic vents and thereby accessible to our observation. However, much magma solidifies below the surface and can only be observed as the end product, an igneous rock, from which the original nature of the magma must be inferred. Magma has been defined as molten rock material, but this definition is not entirely satisfactory. It fails to bring out that magma contains volatile components which are lost as it solidifies but which nevertheless play a significant part in determining the course of crystallization. The definition also fails to emphasize that the solidification of a magma does not take place at a definite temperature, like, for example, the solidification of molten lead, but is generally drawn out in time and place by fractional crystallization. The end products of such crystallization include not only the igneous rock, but also a gas phase and possibly a watery solution. Hence the history of a magma may be long and complex, and the termination is often not easily defined. When does a magma cease to be a magma? When the first solid phase appears? When a quarter, or half, or three-quarters of it has solidified? When all that remains is a watery solution? The diversity of answers to this question is in part responsible for the vigorous controversies regarding the origin of some plutonic rocks, such as granites and granodiorites. Magma is characterized (a) by composition, in that it is predominantly silicate; (b) by temperature, in that it is hot (although the range in temperature may be great, say from 500 to 1200°); and (c) by mobility, in that it will flow. Although a magma is fluid, only a minor part of it need be liquid. In this book the term magma is used to include all naturally occurring mobile rock matter that consists in noteworthy part of a liquid phase having the composition of silicate melt.

The question as to whether or not there is a "primary" magma is one that has been extensively discussed. Many years ago Bunsen decided that there were two primary magmas, granitic and basaltic, and that igneous rocks of other compositions were mixtures of these two types. Bunsen's contention is apparently supported by the abundances of different rock types. The igneous rocks of the globe belong chiefly to two types: granite and basalt. Daly estimated that granites and granodiorites together comprise at least 95% of all intrusives and that basalts and pyroxene andesites make up about 98% of all extrusives (the latter figure seems excessive, since silica-rich extrusives—rhyolites and dacites—are fairly abundant). These relations are clearly of fundamental significance in a consideration of magmatism. Material of basaltic composition may well be the single primary magma from which most igneous rocks have been derived. Major arguments in favor of this are (a) throughout geological time magma of this composition has broken through the crust and poured out on the surface in great floods; (b) the rocks of the great ocean basins are almost entirely basalts; (c) experimental work has shown that differentiation of basaltic magma can yield rocks of widely varying chemical and mineralogical composition, and such differentiation has been observed in natural occurrences. The assumption of a primary basaltic magma in no way implies that magmas of other compositions are nonexistent. This is clearly evident from the wide variety of lavas expelled by volcanoes. The formation of magmas of different compositions can be explained by fractional crystallization of a basaltic magma and the separation of residual magmas of different types, by assimilation of material of different composition, and by the escape of volatile substances, all of which either singly or together are adequate to produce significant changes in composition.

THE CHEMICAL COMPOSITION OF MAGMAS AND IGNEOUS ROCKS

Clarke and Washington calculated the average composition of igneous rocks, with the following results:

SiO_2	Al_2O_3	Fe_2O_3	FeO	MgO	CaO	Na_2O	K_2O	H_2O	TiO_2
59.14	15.34	3.08	3.80	3.49	5.08	3.84	3.13	1.15	1.05

All others less than 0.30 each

It must, of course, be borne in mind that these figures are the average of several thousand analyses and are of no greater or less significance than is implied in that statement; they do not represent the composition of a primary magma from which all igneous rocks may be derived, nor the composition of any particular magma; minor constituents are omitted, as are the volatiles known to be present in all magmas. Nevertheless, the figures do indicate that the following elements predominate: O, Si, Al, Fe, Mg, Ca, Na, and K; that is, magmas are multicomponent systems of these and other elements.

The different components do not vary randomly but are interdependent. The frequency relations of the commonest rock-forming oxides have been worked out on the basis of Washington's collection of 5159 superior analyses of fresh rocks (Figures 5.1
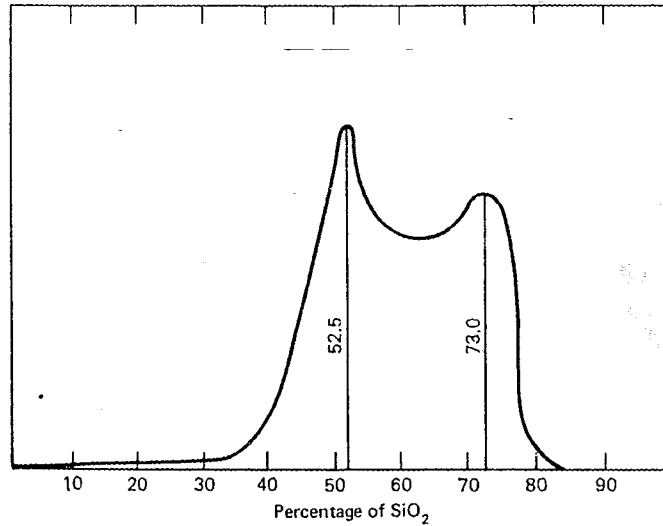
Figure 5.1 The frequency distribution of silica percentage in
analyses of igneous rocks.
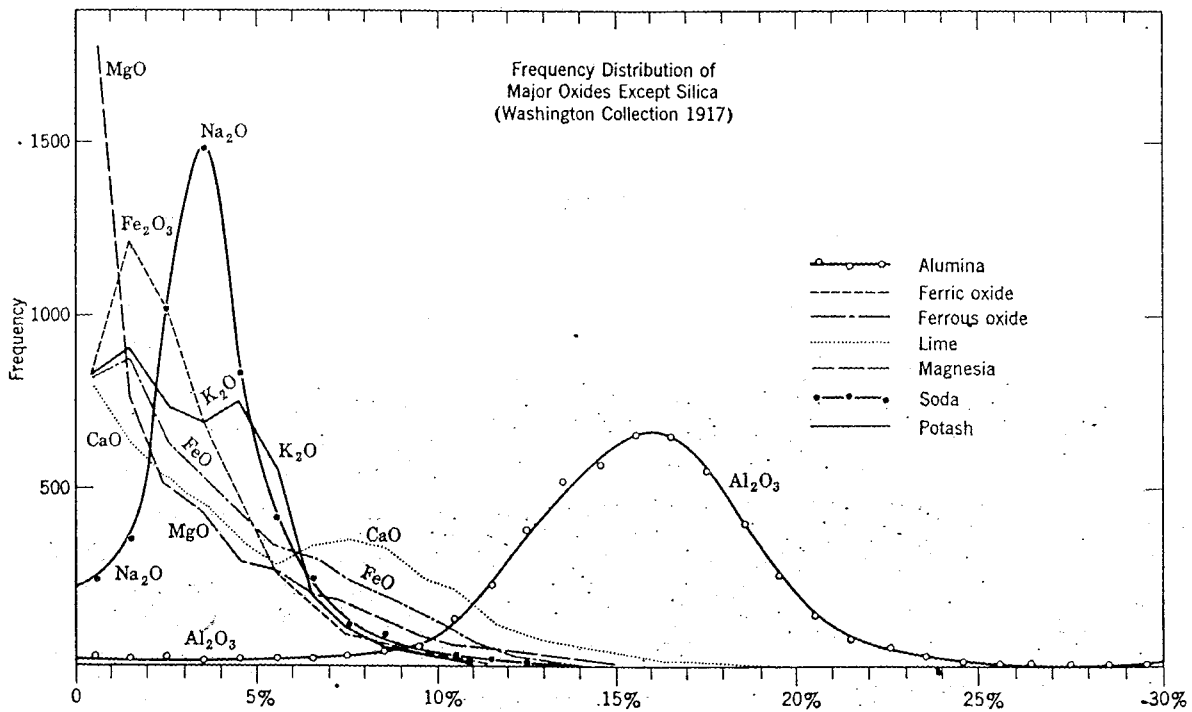
Figure 5.2 The frequency distribution of the percentages of the major oxides in analyses of igneous rocks. (Richard-
son and Sneesby, *Mineralog. Mag.* **19**, 309, 1922)

and 5.2). The dominant oxide is silica; most igneous rocks contain between 30 and 80% SiO_2. Its variation is not at all regular, as is clearly seen from Figure 5.1, there being two frequency maxima, at 52.5% SiO_2 and 73.0% SiO_2; i.e., rocks with about 52.5% SiO_2 and 73.0% SiO_2 are the commonest. This corresponds with field experience that the most abundant igneous rocks are granite and basalt. Alumina varies commonly between 10 and 20%. Low Al_2O_3 is characteristic of rocks with little feldspar or feldspathoid, i.e., ultrabasic types; high Al_2O_3 is characteristic of anorthosites and rocks containing much nepheline. Soda shows a very symmetrical variation, the common range being from 2 to 5%; Na_2O rarely exceeds 15%. The curve for K_2O is less regular, but most rocks have less than 6% K_2O, and only rarely does it exceed 10%. The curves for FeO and Fe_2O_3 are similar; both show a frequency peak at about 1.5% and fall off more or less regularly toward higher percentages. The sum of iron oxides in igneous rocks seldom exceeds 15%, except in magmatic iron ores. The frequency curve for magnesia percentage is very asymmetrical; most rocks have a low MgO content, and only ultrabasic types rich in pyroxene and/or olivine have more than 20%. The curve for CaO is similar to that for MgO but is biomodal; most rocks have less than 5% CaO, but a peak at about 8% is caused by the CaO contents of basaltic rocks. Water may reach 10% in a few volcanic glasses, but generally an igneous rock containing more than 2% H_2O has acquired the excess by alteration. Three minor constituents that should be determined in a good rock analysis are TiO_2, P_2O_5, and MnO, all of which are present in most igneous rocks. Many other elements may be found in small amounts; their occurrence is discussed later in this chapter.

THE MINERALOGICAL COMPOSITION OF IGNEOUS ROCKS

Although more than 1000 different minerals are known, the number of species present in more than 99% of the igneous rocks is very small. Apart from the seven principal minerals or mineral groups (the silica minerals, feldspars, feldspathoids, olivine, pyroxenes, amphiboles, and micas), only magnetite, ilmenite, and apatite are commonly found, and then usually in very small amounts. A statistical study of about 700 petrographically described igneous rocks gave the following average mineralogical composition: quartz 12.0%, feldspars, 59.5%, pyroxene and hornblende 16.8%, biotite 3.8%, titanium minerals 1.5%, apatite 0.6%, other accessory minerals 5.8%. It is interesting to compare this average *mode* with the *norm* calculated from the average composition of igneous rocks. This is Q 10.02, or 18.35, ab 32.49, an 15.29, di 6.45, hy 8.64, mt 4.41, il 1.98, ap 0.67.

The mineralogical composition is a basic criterion in the classification of igneous rocks. Several hundred rock types have been named, but of these only a few are of common occurrence. Figure 5.3 is a diagrammatic representation of the common types, showing the mineralogical composition in terms of the important rock-forming minerals. The diagram serves to emphasize that rock types grade into each other, and a rock name is a convenient pigeonhole rather than a species of fixed composition.

The principal minerals and mineral groups are now discussed in detail.
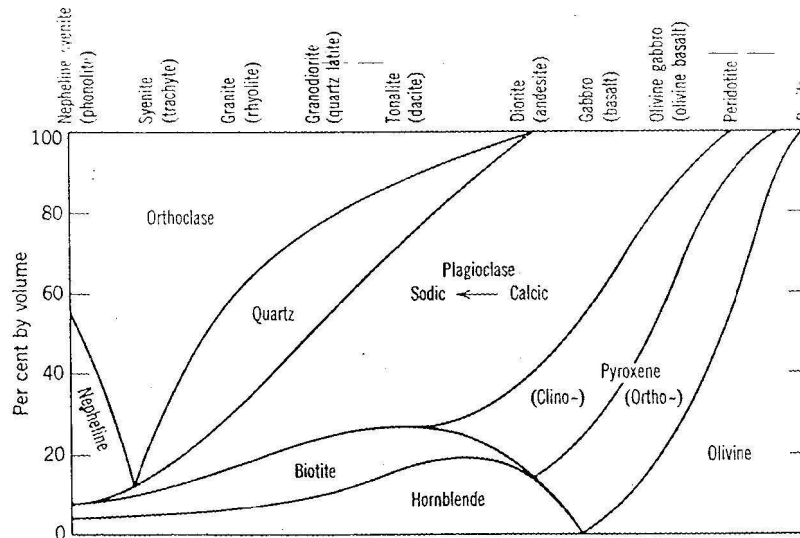
Figure 5.3   Approximate mineralogical composition of the commoner types of
igneous rocks (effusive types in brackets).

The Silica Minerals

Silica occurs in nature as seven distinct minerals: quartz (including chalcedony), tridymite, cristobalite, opal, lechatelierite, coesite, and stishovite. Of these, quartz is very common; tridymite and cristobalite are widely distributed in volcanic rocks and can hardly be called rare; opal is not uncommon; lechatelierite (silica glass) is very rare. Coesite and stishovite are high-pressure forms first made in the laboratory and later recognized in the shocked sandstone at Meteor Crater in Arizona, where they were evidently formed by the instantaneous high pressure of the meteorite impact. The high density of 4.29 for stishovite is due to a change from four- to six-coordination. Although the Si—O bonds in the six-coordinated octahedral structure are longer than in the four-coordinated tetrahedral structure, the closer packing of the oxygens produces a great increase in density over the other silica polymorphs.

Quartz, cristobalite, and tridymite are the forms of silica that may occur in igneous rocks, and their relationships will be discussed in detail. These three forms illustrate the phenomenon of enantiotropism. Each has its own stability field; at atmospheric pressure quartz is the stable form up to 867°, tridymite between 867° and 1470°, and cristobalite from 1470° up to the melting point at 1713°, from 1713° to the boiling point liquid silica is the stable phase. The one-component system SiO_2 has been extensively studied under varying conditions of temperature and pressure; the equilibrium diagram is given in Figure 5.4.

The addition of small amounts of H_2O to this system produces remarkable results. It has been investigated at temperatures up to 1300° and pressures of water vapor

up to 2000 kg/cm². The results are indicated with dotted lines on Figure 5.4. The liquidus is lowered by several hundred degrees. Thus, at pressures above 1400 kg/cm², quartz melts at about 1125°, a depression of some 600° in the melting point; the liquid so formed is a hydrous silica melt containing about 2.3% H_2O. The stability field of tridymite is greatly contracted, and, whereas in the dry system tridymite has no stable melting point, in the presence of water vapor it melts to a hydrous liquid at pressures above 400 kg/cm². Water vapor under pressure thus has a tremendous fluxing power for SiO_2 and clearly will greatly influence the transportation of silica in a fluid form. The large effects produced by a small weight percent of water are a reflection of its low molecular weight, which results in its mole fraction in solution being comparatively large.

These three polymorphs of silica are all built of tetrahedral groups of four oxygen atoms surrounding a central silicon atom. The silicon-oxygen tetrahedra are linked
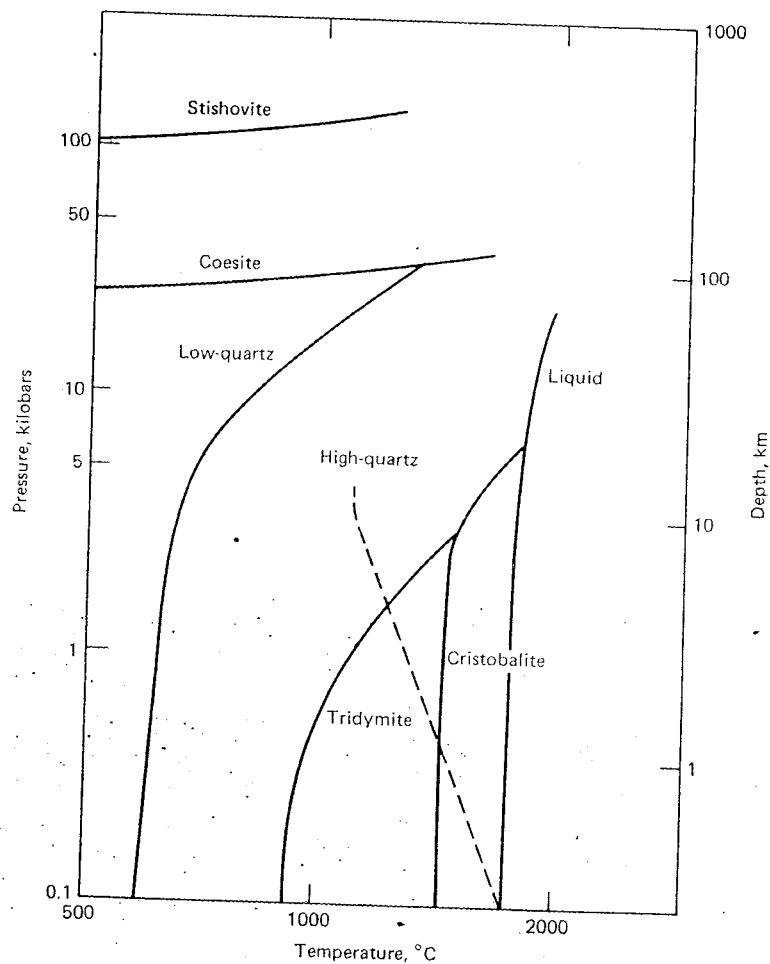
Figure 5.4 Stability fields of silica phases.

together to form a three-dimensional network, but the pattern of linkage is different for each of the three forms; hence the difference in their crystal structures and their properties. Cristobalite and tridymite have comparatively open structures, whereas the atoms in quartz are more closely packed. This is reflected in the densities and refractive indices, which are much lower for cristobalite and tridymite than for quartz:

	Density	Mean Refractive Index
Quartz	2.65	1.55
Cristobalite	2.32	1.49
Tridymite	2.26	1.47
Lechatelierite	2.20	1.46

Each of these three polymorphs has a high- and low-temperature modification. In quartz, for example, the change from the one to the other takes place at 573° at atmospheric pressure. Similarly, high-tridymite changes into low-tridymite between 120 and 160° and high-cristobalite into low-cristobalite between 200 and 275°. The inversion of high-and low-temperature forms of the individual species is of quite another order to that between the species themselves. The three minerals have the SiO_4 tetrahedra linked together according to different schemes, and this linkage has to be completely broken down and rearranged for the transformation of one to another. On the other hand, the change from a high-temperature to a low-temperature form does not alter the way in which the tetrahedra are linked. They undergo a displacement and rotation which alters the symmetry of the structure without breaking any links. The high-temperature modification is always more symmetrical than the low-temperature modification.

The high-low transformation of each mineral takes place rapidly at the transition temperature and is reversible. The changes from one polymorphic form to another are extremely slow and sluggish, and the existence of tridymite and cristobalite as minerals shows that they can remain unchanged indefinitely at ordinary temperatures. Once formed, the type of linking in tridymite and cristobalite is too firm to be easily broken, and it is possible to study the high-low inversions of tridymite and cristobalite at temperatures at which they are really metastable forms. As pointed out in Chapter 4, the presence of foreign elements in the structure may have a stabilizing effect on tridymite and cristobalite. The few comprehensive analyses of these minerals show the presence of Na and Al, suggesting a substitution of NaAl for Si in the open structures; quartz on the other hand is generally very pure SiO_2. Two other phenomena of great significance should be mentioned here:

1. Even at temperatures below 867°, especially when crystallization takes place rapidly (for example, in the presence of mineralizers, such as hot gases), cristobalite and/or tridymite may crystallize, although quartz is the stable phase.
2. High-quartz and low-quartz are formed only within their stability fields, never at higher temperatures.

From these facts the following conclusions can be drawn: quartz in an igneous rock signifies that its crystallization from the magma took place below 867° (with due regard for the effect of pressure); the presence of cristobalite or tridymite, on the other hand, proves nothing as to the temperature of crystallization.

As already pointed out, at ordinary temperatures quartz is always present as low-quartz. By the crystal form, nature of the twinning, and other properties of less diagnostic importance, it is often possible to determine the original form. In this way it has been shown that in nearly all the quartz-bearing igneous rocks this mineral crystallized originally as high-quartz, that is, above 573°. In quartz veins and some pegmatites it crystallized originally as low-quartz. It may therefore be concluded that the crystallization of the magma corresponding to the commonest quartz-bearing rocks took place above 573° and the residual crystallization in part at least at lower temperatures.

The Feldspar Group

plagioclase (Triclinic)
Alkali (monoclinic)

The feldspars owe their importance to the fact that they are the most common of all minerals. They are closely related in form and physical properties, but they fall into two groups: the potassium and barium feldspars, which are monoclinic or very nearly monoclinic in symmetry, and the sodium and calcium feldspars (the plagioclases), which are definitely triclinic. An ammonium feldspar, buddingtonite, also monoclinic, was described in 1964 from a hot-spring deposit in California. A point of great interest is the solid solution between albite, $NaAlSi_3O_8$, and anorthite, $CaAl_2Si_2O_8$. The theory that feldspars of intermediate composition were mixed crystals of these two components was proposed by Tschermak in 1869. It is now known that NaSi often sustitutes for CaAl, but Tschermak's theory is of historic importance as a first suggestion that so radical a substitution is possible.

The general formula for the feldspars can be written WZ_4O_8, in which W may be Na, K, Ca, and Ba, and Z is Si and Al, the Si:Al ratio varying from 3:1 to 1:1. Since all feldspars contain a certain minimum amount of Al, the general formula may be somewhat more specifically stated as $WAl(Al,Si)Si_2O_8$, the variable (Al,Si) being balanced by variation in the proportions of univalent and bivalent cations.

The structure of the feldspars is a continuous three-dimensional network of SiO_4 and AlO_4 tetrahedra, with the positively charged sodium, potassium, ealcium, and barium situated in the interstices of the negatively charged network. The network of SiO_4 and AlO_4 tetrahedra is elastic to some degree and can adjust itself to the sizes of the cations; when the cations are relatively large (K,Ba) the symmetry is monoclinic or pseudomonoclinic; with the smaller cations (Na,Ca) the structure is slightly distorted and the symmetry becomes triclinic.

The barium-containing feldspars are rare and of no importance as rock-forming minerals, and so we will omit them from further consideration. The feldspars may then be discussed as a three-component system, the components being $KAlSi_3O_8$ (Or), $NaAlSi_3O_8$ (Ab), and $CaAl_2Si_2O_8$ (An). Complexities are introduced by the solid solution relations existing among these three components and the occurrence of polymorphic forms.

The potash-feldspar minerals occur in several distinct forms having different but intergradational optical and physical properties. Sanidine, the monoclinic high-temperature polymorph, occurs in volcanic rocks. Common orthoclase, another monoclinic variety, and microcline (triclinic) are found in a wide variety of igneous and metamorphic rocks which have crystallized at intermediate to low temperatures. Adularia is the name given to a form (which may be either monoclinic or triclinic) with a distinctive crystal habit found in low-temperature hydrothermal veins.

Recent research has clarified the relationship between these forms. Microcline and sandidine are polymorphs with an order-disorder relationship, the Si and Al atoms being randomly distributed over their lattice positions in sanidine but ordered in microcline. The disordered form is the more stable polymorph above about 700°, and microcline has been transformed into sanidine by hydrothermal treatment at this temperature; the reverse transformation has not been achieved in the laboratory, evidently because of the high activation energy required for the ordering of the Si and Al atoms. Orthoclase and adularia are structurally intermediate between sanidine and microcline. Much orthoclase probably crystallized originally as sanidine. Adularia is evidently a metastable form that develops under conditions of rapid crystallization within the stability field of microcline; the rapid crystallization prevents the attainment of an ordered arrangement of Si and Al.

At high temperatures complete solid solution exists between $KAlSi_3O_8$ and $NaAlSi_3O_8$. The more potassic members of the series are monoclinic and are called soda-orthoclase; indeed, most of the orthoclase we identify in rocks is really soda-orthoclase with considerably more potassium than sodium. The more sodic members of the series with greater than 63% albite are triclinic and are called anorthoclase. At lower temperatures solid solutions intermediate between orthoclase and albite are metastable and under conditions of slow cooling break down into an oriented intergrowth of subparallel lamellae, alternately sodic and potassic in composition. Such an intergrown is called perthite or antiperthite. In the perthites the plagioclase occurs as uniformly oriented films, veins, and patches within the orthoclase (or microcline); in the antiperthites this relation is reversed. Many apparently homogeneous specimens of alkali feldspar prove to be perthitic on X-ray examination, the intergrowth being submicroscopic. Perthite, when heated for a long time at 1000°, becomes homogeneous once more. Not all perthites have been formed by exsolution; some are undoubtedly the product of partial metasomatic replacement of an originally homogeneous potash feldspar by sodium-bearing solutions.

X-ray examination of potash feldspar and of albite provides the following explanation of the perthite lamellar intergrowth. The framework of linked SiO_4 and AlO_4 tetrahedra, being similar for the monoclinic and triclinic forms, is continuous throughout the structure. At high temperatures the K and Na ions are randomly distributed in the framework, producing a homogeneous crystal. At lower temperatures ordering may occur with the formation of potassium-rich and sodium-rich lamellae, producing alternate sheets with monoclinic or pseudo-monoclinic and triclinic symmetry, respectively. The a repeat of potash feldspar and albite is markedly different (8.45 Å and 8.14 Å), whereas the b and c repeats are almost identical (12.90 Å, 7.14 Å and 12.86 Å, 7.17 Å). This accounts for the lamellae being approx-

imately parallel to 100, since the b and c repeats coincide in the 100 plane, whereas the a repeat running through the lamellae shortens in the albite regions and lengthens in the potash feldspar regions.

The Or—Ab system has been carefully studied by many investigators (Figure 5.5). Great difficulty has been experienced in the laboratory investigation of this system on account of the viscosity of the melts and their extremely sluggish crystallization. Considerable advances have been made by the addition of water under pressure, which promotes crystallization and lowers its temperature without affecting the general equilibrium relations. A pressure of 2000 kg/cm² of water lowers the crystallization temperature by as much as 300°, thereby confirming the opinion long held
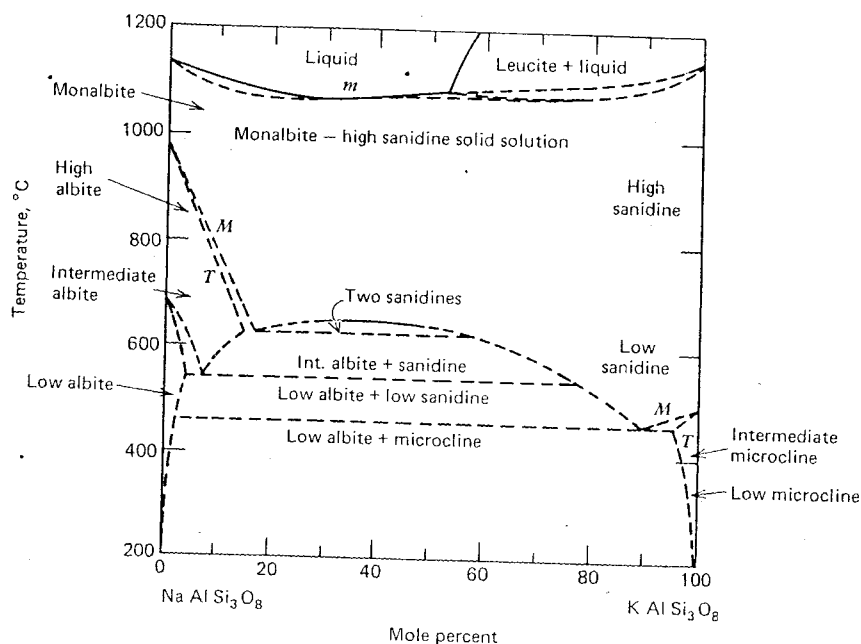
Figure 5.5 Schematic temperature-composition diagram for the $NaAlSi_3O_8$ (Ab)–$KAlSi_3O_8$(Or) join at low pressure to illustrate possible relations for true equilibrium and to show the nomenclature of the phases. The melting relations show a minimum at m and incongruent melting to leucite plus siliceous liquid for potassic compositions. Crystallization from a liquid yields a single feldspar in the monoclinic series monalbite to high sanidine. At low temperature, low albite and low microcline show little solid solution. The top of the unmixing solvus bounds the field of two sanidines. The ordering inversions of the endmember compositions are assumed to be first-order. The solid lines show stable equilibria determined in the laboratory, while the dashed lines show phase boundaries whose nature or exact position is in doubt. M and T indicate monoclinic and triclinic structures. (From J. V. Smith, Feldspar Minerals, Vol. 1, 1974. Springer-Verlag, New York. Used by permission)

among some petrologists as to the importance of the water content in reducing the temperature of magmatic crystallization. The system Or—Ab is not truly binary because of the incongruent melting of Or to give leucite; however, the field of leucite almost disappears at 2000 kg/cm² pressure.

The shape and position of the solidus curve are not significantly altered by the presence or absence of water, since the equilibrium is one between solids, in which water does not participate in any of the phases. This curve, dividing the two-feldspar field (i.e., the perthite field) from the one–feldspar field, has a maximum at about 680°C at a composition close to 55% Ab. Thus crystallization above 680° in the Or—Ab system gives a single feldspar at any composition. Any point on the solvus represents the minimum temperature of stable existence for homogeneous feldspar of that composition; if equilibrium is maintained below that temperature, the feldspar will begin to unmix.

The system Ab—An, long quoted as a classic example of perfect solid solution, also shows unexpected complications (Figure 5.6). The solidus-liquidus relations, one of the first fruits of research at the Geophysical Laboratory, remain largely unamended. The virtually complete solid solution at high temperatures, however, is affected by the inversion of albite to a low-temperature form at about 700°; the relationship between the low and high forms is evidently an order-disorder one, similar to that between microcline and sanidine. The unit cell of anorthite has a doubled c axis with respect to that of albite, and two distinct structures are found, primitive (P) and body-centered (I) respectively. These complications were first suspected
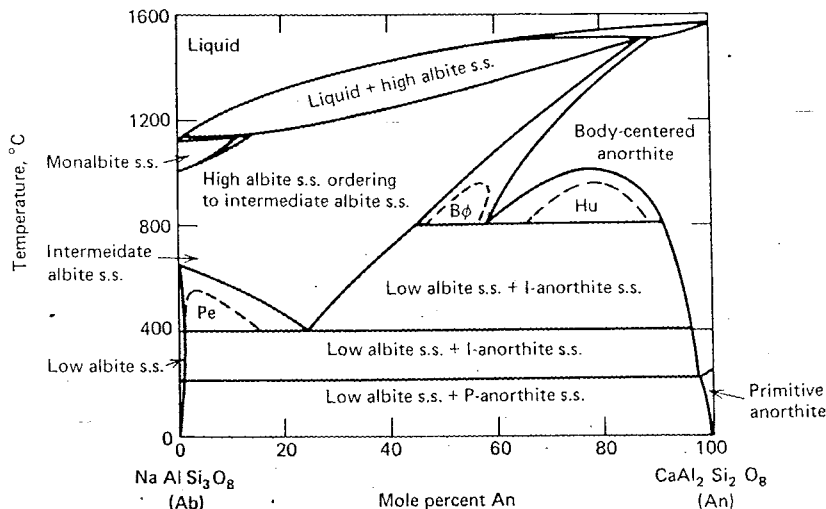
Figure 5.6 Possible phase relations in the plagioclase series. The dashed lines shown metastable curves that govern the peristerite, Bøggild, and Huttenlocher intergrowths. All curves are uncertain except for the solidus, liquidus, and the inversion from monalbite to high albite. (From J. V. Smith, *Feldspar minerals,* vol. 1, 1974. Springer-Verlag, New York. Used by permission)

when discrepancies were observed between the optical properties of plagioclases of the same composition from volcanic and plutonic rocks, indicating distinct series of high-temperature and low-temperature plagioclases. The low-albite structure can accept only a little calcium replacing sodium, and in the approximate range An_3 to An_{23} plagioclases are usually submicroscopic intergrowths, known as peristerite, of sodium-rich and calcium-rich regions. Plagioclases more calcic than about An_{23} crystallize in variants of the anorthite structure, and submicroscopic intergrowths of sodium-rich and calcium-rich regions may be present also (Bøggild and Huttenlocher intergrowths). Thus many plagioclases that appear optically homogeneous prove to consist of complex intergrowths when examined by X-ray techniques.

The mutual miscibility of Or and An is nearly zero at all temperatures, hence varieties intermediate between these two components do not occur. Careful analyses of homogeneous plagioclases have shown that the potassium content is usually very low, indicating that the normal amount of the $KAlSi_3O_8$ component in solid solution is about 1% and that it seldom rises above 5%.

The composition of any feldspar is conveniently expressed by the use of the symbols Or, Ab, and An for the pure components: thus Ab_{32} and An_{68} would be a plagioclase falling in the labradorite section of the series, and $Or_{26}Ab_{66}An_8$ is the composition of a possible anorthoclase.

The Feldspathoids

The feldspathoids are a group of alkali-aluminum silicates that appear in place of the feldspars when an alkali-rich magma is deficient in silica. They are never associated with primary quartz. The following minerals are the more important members of this group:

Leucite	$KAlSi_2O_6$
Kaliophilite	$KAlSiO_4$
Kalsilite	$KAlSiO_4$
Nepheline	$NaAlSiO_4$
Sodalite	$Na_8Al_6Si_6O_{24}(Cl_2)$
Nosean	$Na_8Al_6Si_6O_{24}(SO_4)$
Cancrinite	$Na_8Al_6Si_6O_{24}(HCO_3)_2$

Analcime, $NaAlSi_2O_6 \cdot H_2O$, is sometimes included among the feldspathoids; it appears occasionally as a primary mineral of igneous rocks deficient in silica.

The feldspathoids are not a homogeneous series like the feldspars or the pyroxenes. They are grouped together more on petrographical than mineralogical similarities. Structurally, the feldspathoids belong to the tektosilicates, the SiO_4 and AlO_4 tetrahedra being linked as in the feldspars, whereas the metal ions (and the chloride, sulfate, and carbonate ions, when present) fit into cavities in this framework. The feldspathoids are readily attacked by acids. This characteristic is evidently due to the comparatively high Al:Si ratio; the aluminum is removed in solution and the lattice then collapses, often with the formation of gelatinous silica.

Leucite is the commonest potassium feldspathoid. It is abundant in the volcanic

rocks of a few regions. It is not found in plutonic rocks, and laboratory investigations have indicated that leucite is not a stable phase in the K_2O—Al_2O_3—SiO_2 system at high pressures. Analyses of leucite show that a little sodium may replace potassium in the structure. Kaliophilite and kalsilite are mineralogical curiosities; kaliophilite has been recorded in Vesuvian lavas, kalsilite in lavas from East Africa. Their optical properties are so similar to those of nepheline that misidentification could readily occur. On this account they may be of wider distribution than present records indicate.

Nepheline is the commonest of the feldspathoids and is found in both volcanic and plutonic rocks. It is isomorphous with high-tridymite (cf. the formulas $NaAlSiO_4$ and $SiSiO_4$). The analogy between $NaAlSiO_4$ and SiO_2 is further emphasized by the inversion at 1248° of nepheline to carnegieite, a high-temperature form which is isomorphous with high-cristobalite; carnegieite has not been found as a mineral. Analyses of nepheline generally show an excess of Si over the theoretical amount, the ratio Si:Al ranging up to 1.4 and the sodium being correspondingly deficient. The phenomenon is one of Si replacing Al and the consequent omission of sodium ions to preserve electrical neutrality. Nepheline always contains some potassium replacing sodium, the Na:K ratio often being close to 3:1; this reflects the atomic structure, in which one of the four positions that can be filled by an alkali ion is larger than the other three and preferentially accommodates potassium.

The Pyroxene Group

The pyroxenes are a group of minerals closely related in crystallographic and other physical properties, as well as in chemical composition, although they crystallize in two different systems, orthorhombic and monoclinic. The group characteristics are the outward expression of common atomic structure. The tetrahedral SiO_4 groups are linked together into chains by the sharing of one oxygen atom between two adjacent groups; that is, in each group two oxygen atoms also belong half to the groups on each side, giving the Si:O ratio of 1:3. The silicon-oxygen chains lie parallel to the vertical crystallographic axis and are bound together laterally by the metal ions. The distinct prismatic habit of pyroxene crystals is a consequence of this internal structure, as is also the typical cleavage, which takes place between the silicon-oxygen chains.

The chemical composition of the pyroxenes can be expressed by the general formula $(W)_{1-p}(X, Y)_{1+p}Z_2O_6$, in which the symbols W, X, Y, and Z indicate elements having similar ionic radii and capable of replacing each other in the structure. In the pyroxenes these elements may be

W: Na, Ca
X: Mg, Fe^{2+}, Li, Mn
Y: Al, Fe^{3+}, Ti
Z: Si, Al (in minor amount)

The pyroxenes fall naturally into two divisions, those of orthorhombic symmetry and those of monoclinic symmetry.

On the basis of chemical composition and crystal system the following distinct species are recognized:

Orthorhombic Pyroxenes

Enstatite	$MgSiO_3$
Hypersthene	$(Mg,Fe)SiO_3$.

Monoclinic Pyroxenes

Clinoenstatite	$MgSiO_3$
Clinohypersthene	$(Mg,Fe)SiO_3$
Diopside	$CaMgSi_2O_6$
Hedenbergite	$CaFe^{2+}Si_2O_6$
Augite	Intermediate between diopside and hedenbergite with some Al
Pigeonite	Intermediate between augite and clinoenstatite-clinohypersthene
Aegirine (acmite)	$NaFe^{3+}Si_2O_6$
Jadeite	$NaAlSi_2O_6$
Spodumene	$LiAlSi_2O_6$
Johannsenite	$CaMnSi_2O_6$

Spodumene, although a member of the monoclinic pyroxenes, is not closely related to the other species in the group, and no intermediate compounds are known. It is not an important rock-forming mineral, being confined to the complex granite pegmatites. Jadeite is a rare mineral of metamorphic rocks. Johannsenite·is found only as a vein mineral. These species are not considered here.)

The orthorhombic pyroxenes range in composition from pure $MgSiO_3$ to about 90% $FeSiO_3$. The compound $FeSiO_3$ is not a stable phase at high temperatures and low pressure; from a melt of its composition silica and fayalite (Fe_2SiO_4) will crystallize. It has been reported in the monoclinic form from lithophysae of an obsidian, where it would have formed at moderate temperatures. The common orthopyroxenes of igneous rocks are all magnesium-rich. The orthorhombic pyroxenes have been divided into a number of subspecies based on composition, similar to the division of the plagioclase series. The utility of such subdivision is dubious; subspecific names are a burden on the nomenclature and the memory, and in general a designation in the form hypersthene (En_x) is to be preferred, just as plagioclase (An_{25}) is preferable to a less precise identification as oligoclase.

When heated above 985° pure $MgSiO_3$ changes into protoenstatite, another orthorhombic phase with a different structure, which is stable up to 1557°, when it melts incongruently to give forsterite plus liquid. Protoenstatite has not been found as a mineral, but on rapid cooling it changes into clinoenstatite, a monoclinic phase. Hypersthene has a corresponding monoclinic phase known as clinohypersthene. Clinoenstatite and clinohypersthene are almost unknown in terrestrial rocks but have been recognized in some meteorites. Intermediate compounds between these substances and augite are, however, found in basic volcanic and hypabyssal rocks and

are known as pigeonite. Analyses of pigeonite suggest a maximum of about 10 atom % Ca in $W + X + Y$. The diopside-hedenbergite series is well established throughout the whole range in composition. A limited amount of aluminum can enter into this series, and such aluminous pyroxenes are called augite. The introduction of Al into a diopside means that this Al must be divided between the Y and the Z lattice positions, otherwise the valence demands would not be satisfied. Aegirine is connected with the diopside-hedenbergite series and augite by members of intermediate composition, which are grouped under the name aegirine-augite.

The compositions of the naturally occurring pyroxenes are extremely variable because of the possibilities of atomic substitution. In the general formula p is zero or near zero in the diopside-hedenbergite and aegirine-jadeite series and 1 or close to 1 in the orthorhombic pyroxenes, their monoclinic dimorphs, and in the pigeonites. Magnesium and ferrous iron are completely interchangeable. Reliable analyses of the common pyroxenes, the diopside-hedenbergite-augite group, show a maximum of about 7% Fe_2O_3 (corresponding to 10% Fe^{3+} in Y); about 8% Al_2O_3 (corresponding to 10% or a little more Al replacing Si and a smaller Al replacement in Y); and about 1.5% TiO_2. Minor constituents recorded in pyroxene analyses include chromium (up to 1.2% Cr_2O_3 in some diopsides and augites), vanadium (4% V_2O_3 in aegirine from Libby, Montana), and manganese. The miscibility gap between augite and $(Mg,Fe)SiO_3$ is considerably less at high than at moderate temperatures; in volcanic rocks the pyroxenes may be low-calcium augite (sometimes termed subcalcic augite) and/or pigeonite, whereas in plutonic rocks the pyroxenes are normal augites and/or orthorhombic pyroxenes almost free of calcium

The Amphibole Group

The amphibole group comprises a number of species, which, although falling both in the orthorhombic and monoclinic systems, are closely related in crystallographic and other physical properties, as well as in chemical composition. They form isomorphous series, and extensive replacement of one ion by others of similar size can take place, giving rise to very complex chemical compositions. The species of the amphibole group form a series parallel to those of the allied pyroxene group; they were originally looked upon as complex metasilicates dimorphous with the corresponding pyroxenes. This is not so, however; the amphiboles contain essential (OH) groups in their structure, and the Si:O ratio is 4:11, not 1:3 as in the pyroxenes. The true nature of the amphiboles has been elucidated by X-ray studies, which have shown that the fundamental unit in their structure is a double chain of linked silicon-oxygen tetrahedra; in effect it is two single chains with alternate tetrahedra linked by the sharing of the oxygen, giving an Si:O ratio of 4:11, instead of 1:3 as in single chains. In the structure these double chains lie parallel to the vertical crystallographic axis and are bound together laterally by the metal ions. The binding force between the chains is not as strong as the Si-O bonds along the chain. This is reflected in the well-known fibrous or prismatic nature of the amphiboles parallel to the c axis and in the good prismatic cleavage. Another interesting feature of the amphibole structure is the

presence of open spaces within the lattice into which an extra alkali atom for each two Si_4O_{11} groups will fit.

The difference in chemical composition between compounds of the amphibole type and corresponding compounds of the pyroxene type is not great, and it is understandable that it was overlooked for many years and that the amphiboles and pyroxenes of similar composition were considered to be polymorphs of the same substance. This is illustrated by the theoretical compositions of $MgSiO_3$ and $Mg_7(Si_4O_{11})_2(OH)_2$

	MgO	SiO_2	H_2O
$MgSiO_3$	40.0	60.0	—
$Mg_7(Si_4O_{11})_2(OH)_2$	36.2	61.5	2.3

A general formula for all members of the amphibole group can be written $(WXY)_{7-8}(Z_4O_{11})_2(O,OH,F)_2$, in which the symbols W,X,Y,Z indicate elements having similar ionic radii and capable of replacing each other in the structure: W stands for the large metallic cations Ca and Na (K is sometimes present in small amounts); X stands for the smaller metallic cations Mg and Fe^{2+} (sometimes Mn); Y for Ti, Al, and Fe^{3+}; and Z for Si and Al. In the general formula the degree of atomic substitution is as follows:

1. Al may replace Si in the Si_4O_{11} chains to the extent of $AlSi_3O_{11}$ (the amount of replacement is a function of the conditions of formation; high-temperature amphiboles can be more aluminous than low-temperature amphiboles).
2. Fe^{2+} and Mg are completely interchangeable.
3. The total (Ca, Na, K) may be zero or near zero or may vary from 2 to 3; however, total Ca never exceeds 2, and K is present only in minor amounts.
4. OH and F are completely interchangeable. The maximum is 2, but it may be less, presumably by O replacing OH or F.

With these possibilities it is clear that the composition of the amphiboles may be very complex. However, on the basis of chemical composition and crystal structure five distinct series are recognized.

Orthorhombic

Anthophyllite series $(Mg,Fe)_7(Si_4O_{11})_2(OH)_2$ (Mg predominant over Fe)

Monoclinic

Cummingtonite series $(Fe,Mg)_7(Si_4O_{11})_2(OH)_2$ (Fe predominant over Mg)
Tremolite series $Ca_2(Mg,Fe)_5(Si_4O_{11})_2(OH)_2$
Hornblende series $Ca_2Na_{0-1}(Mg,Fe,Al)_5[(Al,Si)_4)_{11}]_2(OH)_2$
Alkali amphibole series   (Na > Ca), e.g.,
     Glaucophane $Na_2Mg_3Al_2(Si_4O_{11})_2(OH)_2$
     Riebeckite $Na_2Fe_3^{2+}Fe_2^{3+}(Si_4O_{11})_2(OH)_2$
     Arfvedsonite $Na_3Fe_4^{2+}Fe^{3+}(Si_4O_{11})_2(OH)_2$

The anthophyllite series corresponds to the orthorhombic enstatite-hypersthene series in the pyroxene group. The ratio Mg:Fe ranges from 7:0 up to about 1:1; Al

is often present and can replace Si in (Si_4O_{11}) up to $(AlSi_3O_{11})$, with corresponding replacements in the (WXY) elements. Other substitutions are minor. Members of the anthophyllite series have been found only in metamorphic rocks.

The cummingtonite series corresponds to the clinohypersthene series in the pyroxene group. The ratio Fe:Mg ranges from 7:0 to about 1:2; that is, the cummingtonite series and the anthophyllite series overlap in the middle composition range. Rocks containing these two amphiboles side by side have been described. Members of the cummingtonite series occur mainly in metamorphic rocks.

The tremolite series corresponds to the diopside-hedenbergite series in the pyroxene group. In tremolite magnesium is replaceable by ferrous iron and also in part by aluminum and ferric iron, silicon in part by aluminum; titanium and fluorine may be present, and an additional sodium ion for each two (Si_4O_{11}) groups may enter the structure. The product of all these substitutions is the hornblende series, analogous to augite in the pyroxene group but considerably more complex. Thus the mineral known as hornblende has a very wide range of composition and a correspondingly wide range in optical properties. Most hornblende is green, but there is a dark brown variety, long known as basaltic hornblende, which is sometimes considered as a separate series on account of its distinctive properties.

The alkali amphiboles can be considered as derived from the hornblende series by the partial or complete substitution of Na for Ca. The best known of these soda amphiboles are glaucophane and the related riebeckite and arfvedsonite, although other varieties have been described.

In view of the similarity in composition between corresponding amphiboles and pyroxenes the circumstances under which one or the other or both will crystallize from a magma have been the subject of considerable speculation. Amphiboles occur more often in plutonic rocks than in volcanic rocks, evidently because the incorporation of OH groups in the structure is favored by crystallization under pressure. However, the composition of the magma may also be significant. The Ca:Fe:Mg ratios of igneous hornblendes and pyroxenes show characteristic differences; many hornblendes fall in the composition gap between augite and the pigeonites and orthorhombic pyroxenes.

The Olivine Group

The minerals of the olivine group are silicates of bivalent metals and crystallize in the orthorhombic system. There are a number of species:

Forsterite	Mg_2SiO_4
Fayalite	Fe_2SiO_4
Olivine	$(Mg,Fe)_2SiO_4$
Tephroite	Mn_2SiO_4
Monticellite	$CaMgSiO_4$
Kirschsteinite	$CaFeSiO_4$
Glaucochroite	$CaMnSiO_4$

The only common rock-forming olivines are the magnesium-iron compounds, although monticellite is found in metamorphosed limestones and monticellite and kirschsteinite have been recorded from basic igneous rocks. The other species are known from ore deposits in metamorphosed limestones.

X-ray studies on olivine show that it is built up of independent SiO_4 tetrahedra and that the magnesium and ferrous ions lie between irregular groups of six oxygens and belong to two sets that are not structurally identical. In keeping with this structural type, which does not have extended chains or rings of SiO_4 tetrahedra, the minerals of the olivine group have no tendency to form fibrous or platy crystals and generally occur as equidimensional crystals. The close-packed structure of olivine is reflected in comparatively high density (forsterite, 3.22; cf. enstatite, 3.18; anthophyllite, 2.96; talc, 2.82) and high refractive indices.

The composition of olivine generally corresponds closely to $(Mg,Fe)_2SiO_4$, there being little replacement by other elements. Substitution by calcium is evidently strongly temperature-dependent, since olivine from plutonic rocks seldom contains more than 0.1% CaO, whereas that from volcanic rocks usually contains more than this amount, ranging up to a maximum around 1% CaO. Manganese is present in most olivines and correlates positively with Fe content, ranging from about 0.1% in forsterites up to 1 to 2.5% in fayalite. Olivines from dunites generally contain some nickel, often about 0.3%. A noteworthy feature of olivine is the practical absence of aluminum; evidently replacement of Mg and Si by Al does not occur in the olivine structure.

The Mica Group

With their characteristic perfect basal cleavage the members of the mica group are readily recognizable. The composition of individual specimens may be very complex, but a general formula of the type $W(X,Y)_{2-3}Z_4O_{10}(OH,F)_2$ can be written for the group as a whole. In this formula W is generally potassium (Na in paragonite); X and Y represent Al, Li, Mg, Fe^{2+}, and Fe^{3+}; Z represents Si and Al, the Si:Al ratio being generally about 3:1. The different species form a typical isomorphous group, but the phase relations have not yet been fully worked out. Two members of the group frequently crystallize together in parallel position. Biotite crystallizes in this way with muscovite, muscovite with lepidolite, and so forth. In the following list the formulas have been simplified to an ideal type conforming to the structure established by X-ray investigation.

Muscovite	$KAl_2(AlSi_3O_{10})(OH)_2$
Paragonite	$NaAl_2(AlSi_3O_{10})(OH)_2$
Phlogopite	$KMg_3(AlSi_3O_{10})(OH)_2$
Biotite	$K(Mg,Fe)_3(AlSi_3O_{10})(OH)_2$
Lepidolite	$KLi_2Al(Si_4O_{10})(OH)_2$

The structural scheme of mica is as follows: The basic units, the SiO_4 tetrahedra, are each linked by three corners to neighboring tetrahedra, forming a sheet. Each

SiO$_4$ tetrahedron thus has three shared and one free oxygen; hence the composition and valency are represented by (Si$_4$O$_{10}$)$^{4-}$. Two of these sheets of linked tetrahedra are placed together with the vertices of the tetrahedra pointing inward. These vertices are cross-linked by Al in muscovite or by Mg and Fe in phlogopite and biotite. Hydroxyl groups are incorporated and linked to Al, Mg, or Fe alone. A firmly bound double sheet is thus produced with the bases of the tetrahedra on each outer side. The structure is a succession of such double sheets, with the potassium ions placed between them (illustrated in Chapter 6, Figure 6.4).

The common mica of igneous rocks is biotite. Muscovite is present in some granites. Lepidolite has been recognized in a few granites, but its typical mode of occurrence is in granite pegmatites. Phlogopite is sometimes found in igneous rocks rich in magnesium and poor in iron, such as peridotites, but is more common in metamorphosed limestones and in some pegmatites. Paragonite is a rare mineral of schists.

The common occurrence of biotite in igneous rocks in contrast to the limitation of muscovite to pegmatites and some granites has been clarified by the work of Yoder and Eugster (Figure 5.7). They found that the stability curve for phlogopite lies about 300° higher than that for muscovite and is well above minimum melting curve for granite. This means that phlogopite (and biotite) can form directly from a magma at normal crystallization temperatures. The stability curve for muscovite, on the other hand, lies below the minimum melting curve for granite at low pressure, intersecting it at about 700° and 1500 atm of water-vapor pressure; hence muscovite
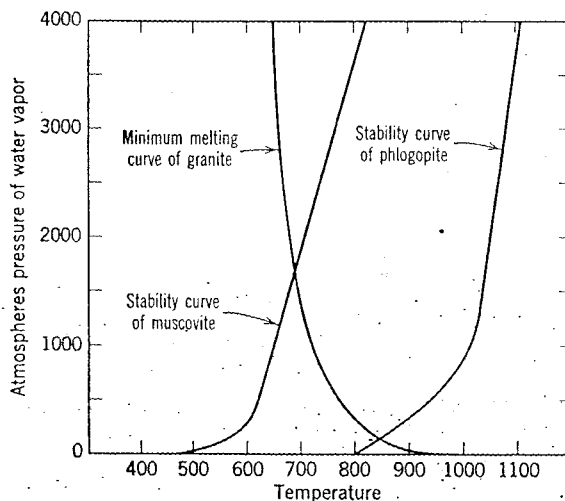
Figure 5.7 Relationship between the stability curves of muscovite and phlogopite and the minimum melting curve of granite. (After Yoder and Eugster, *Geochim. et Cosmochim. Acta* 6, 195, 1954)
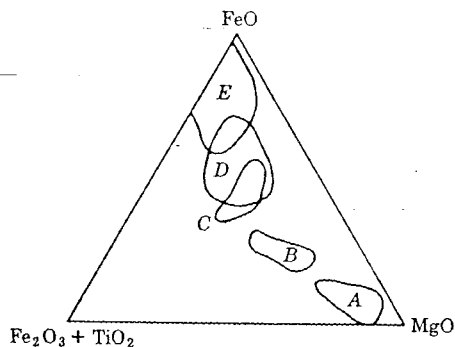
Figure 5.8 Variation of chemical composition of biotite with rock type; (*A*) in ultrabasic rocks; (*B*) in gabbros; (*C*) in diorites; (*D*) in granites; (*E*) in pegmatites. (After Heinrich, *Am. J. Sci.* **244**, 836, 1946)

in a granite implies crystallization at high water-vapor pressure, that is, considerable depth, or later introduction into the already crystallized rock.

Biotites .from igneous rocks are exceedingly variable in chemical composition. Magnesium and ferrous iron are completely diadochic, and all types are known, from iron-free (i.e., phlogopite) to varieties in which nearly all the magnesium is replaced by iron. Ferric iron may replace half or more of the six-coordinated aluminum. Part of the hydroxyl may be replaced by fluorine, although most analyses of igneous biotites show only small amounts of this element. Minor amounts of Mn, Ti, Li, Na, and Ca are recorded in analyses of biotite, and less commonly Ba, Cr, Ni, Rb, and Cs. There is a general trend from magnesium-rich biotites in ultrabasic rocks to iron-rich biotites in granites and nepheline syenites (Figure 5.8). Aluminum is greatest in biotites from granites and pegmatites and is lowest in biotites from ultrabasic rocks; silicon shows an inverse relationship to aluminum.

THE NATURE OF A SILICATE MELT

The entropy of fusion of silicates is not high, and so the atoms or ions in the molten silicate evidently have a degree of order not greatly different from that in the solid. It thus may be inferred that a good deal of the structural arrangement is preserved on melting. Since the silicon-oxygen bonds are much stronger than other links in silicate structures, it is reasonable to assume that these bonds are present to some extent in the liquid also; that is, the anions in the melt are predominantly polymerized silicon-oxygen tetrahedra linked by sharing oxygens into one-, two-, or three-dimensional networks, similar to those in crystalline silicates but more irregular. The high viscosity of silicate melts is evidently due to the presence of these complex silicon-oxygen groupings.

The degree of polymerization of the silicate anions is markedly affected by the ratio of Si (and Al) to O. The linking of the silicon-oxygen tetrahedra will be greatest at Si:O ratio of 1:2 and will decrease as the proportion of Si decreases. This is borne out by the increase in viscosity of silicate melts as the SiO_2 content increases. The presence of other ions is also significant. In alkali silicate melts the viscosity decreases

in the sequence K > Na > Li; evidently sodium is more effective than potassium in disrupting the Si-O bonds and thereby breaking up the polymerized anions into smaller units, and lithium even more so. Probably bivalent cations are even more effective in this respect; iron-rich melts are of notably low viscosity even at high SiO_2 contents. Laboratory experiments have shown, however, that small amounts of H_2O have a particularly remarkable influence in decreasing the viscosity, evidently because of the strong tendency for the reaction $H_2O + O^{2-} = 2(OH)^-$ to take place. The oxygen links between silicon-oxygen tetrahedra are thereby destroyed by the formation of hydroxyl groups, and the polymerized anions broken down into simpler groups. Since the "molecular weight" of OH is low, a small weight percentage of H_2O in a silicate melt is very effective in eliminating links between silicon-oxygen tetrahedra. Silicon and other complex or polymer-forming cations such as aluminum are termed network formers while elements that give up oxygen to the predominantly covalent cation-oxygen complexes are known as network modifiers.

CRYSTALLIZATION IN SILICATE MELTS

Through the study of crystallization in artificial silicate melts of known composition great advances have been made in the understanding of the geochemistry of igneous rocks. The principles of heterogeneous equilibrium govern crystallization from a liquid; they express the conditions under which only one of the possible crystallization products is to be expected and under what conditions more than one may appear simultaneously. The basic relation is, of course, the phase rule, which states that in any system the number of phases (P) plus the number of degrees of freedom (F) are equal to the number of components (C) plus two, or $P + F = C + 2$. The phase rule has been extended in the consideration of geochemical processes in the following way: In any system the maximum number of phases can be reached only when the number of degrees of freedom are at a minimum. This state can be realized by fixing both temperature and pressure. However, it is extremely unlikely that such conditions will occur during magmatic crystallization, since it proceeds as a rule over a great P-T range; that is, these factors remain variable and thus give two degrees of freedom. Under these circumstances the phase rule becomes $P = C$; that is, in a system of n components at arbitrary temperature and pressure no more than n phases (minerals) can be mutually stable. This extension of the phase rule is due to Goldschmidt and is sometimes known as the *mineralogical phase rule.*

Earlier in this chapter it was remarked that despite their wide range in chemical composition igneous rocks have a comparatively simple mineralogy in that probably 99% are made up essentially of seven minerals or mineral groups. This limitation to the number of phases normally formed by the crystallization of a magma is clearly understandable in terms of the phase rule. Actually, considering a magma as an eight-component system of O, Si, Al, Fe, Mg, Ca, Na, and K, it is improbable that as many as eight phases would crystallize from any composition in that system because the individual components are not completely independent, some being capable of replacing each other atom for atom in minerals.

Since 1905 a steady stream of publications dealing with equilibria in successively more complex silicate systems has been issued by the Geophysical Laboratory. Since the results of this work are fully described and discussed in standard texts on petrology, no attempt is made to cover this field here. However, the following systems are selected as illustrating the major principles that have been developed.

The Pyroxene-Plagioclase System

This system is particularly significant in petrogenesis, since compositions within it are reasonably close to basalts and gabbros, which are often essentially pyroxene-plagioclase rocks. We may start with the three-component system diopside (Di)-anorthite (An)-albite (Ab). The phase relations as worked out by Bowen in 1915 are given in Figure 5.9. It is now known that the system is not strictly ternary, because some Al_2O_3 from the plagioclase enters the diopside; however, the phase relations as established by Bowen are still valid. The equilibrium diagram is divided into a plagioclase field and a diopside field by a boundary curve joining the eutectic points in the Di—An and Di—Ab systems. This boundary curve is known either as a *reaction curve* or a *cotectic line;* it marks the bottom of a valley on the liquidus surface, as can be seen from the temperature contours. From any melt with a composition that

Figure 5.9   The system albite-anorthite-diopside. (Bowen, *Am. J. Sci.* **190,** 167, 1915)

places it in the plagioclase field plagioclase crystallizes first, and the composition of the liquid moves in the directions of the reaction curve; when it reaches this curve, diopside begins to crystallize; thereafter diopside and plagioclase crystallize together, and the composition of the melt follows the reaction curve towards the albite–diopside eutectic. Solidification is complete at a point that varies according to the initial composition of the melt and the extent to which the composition of the feldspar changes during crystallization. The course of solidification is similar for a melt with an initial composition placing it in the diopside field, except that diopside is then the first phase to crystallize.

By the addition of $FeSiO_3$ (Fs) as a fourth component, compositions resembling those of basalts and gabbros can be represented. This system has not been completely worked out in the laboratory, but considerable data are available for it, and Barth established a tentative and simplified equilibrium diagram in the form of a tetrahedron in which each corner represents one component (Figure 5.10). In this diagram A represents Ab, B An, C Di, and D Fs; the point p is the Di—Ab eutectic, q the Di—An eutectic, and r and s the (probable) eutectics between Fs and An and Ab. The system consists of two binary solid solution series, the plagioclase series AB and the pyroxene series CD; four binary eutectic systems AC, BC, BD, and AD, with binary eutectics at p, q, r, and s; and four ternary systems each divided into two fields by the reaction curves pq, qr, rs, and sp; the quaternary system itself is divided into two parts by the *reaction surface*.

Figure 5.10   Tetrahedron illustrating the approximate composition of basaltic lavas. The plane *pqrs* represents the boundary surface separating those lavas that precipitate pyroxene first from those that precipitate plagioclase first. (Barth, *Am. J. Sci.* **231**, 331, 1936)

From a melt with any composition between *AB* and the reaction surface the first phase to crystallize is plagioclase. The melt thereby becomes richer in components *C* and *D*, and eventually its composition lies in the reaction surface *pqrs*. When this surface is reached, pyroxene begins to crystallize together with the plagioclase. From then on the composition of the melt lies in the reaction surface and moves down the temperature gradient, and the crystals of both plagioclase and pyroxene react with the melt and change in composition continuously until solidification is complete. The same possibilities for fractional crystallization as in the simpler systems are present. If fractional crystallization takes place, the early-separated crystals will be rich in An and Di and the last-formed crystals richer in the Ab and Fs components.

Barth studied the composition and sequence of crystallization of a number of basalts and showed that the ideas developed from a consideration of the quaternary system discussed above are borne out by the petrographic data. He was able to derive approximate coordinates for the position of the reaction surface by plotting in the tetrahedron the compositions of those basalts that showed evidence of simultaneous crystallization of plagioclase and pyroxene. The points thus obtained defined the reaction surface; he found that they lay almost exactly in one plane and that the position of this plane agreed with the position that would be expected on the basis of the (incomplete) laboratory study of this system.

The experimental study of the crystallization of natural and synthetic melts of basaltic or near-basaltic composition has been greatly extended since the pioneer work of Bowen and Barth. The principles outlined above remain valid, but the detailed information is much more extensive, and possible complications and variations on the simple system have been elucidated. A comprehensive summary is provided by the monograph of Yoder (1976).

The $KAlSiO_4$—$NaAlSiO_4$—SiO_2 System

Compositions within this system are close to some igneous rocks—the granites, syenites, nepheline syenites, and the corresponding rhyolites, trachytes, and phonolites. Investigation of laboratory melts has shown that fractional crystallization of complex silicate melts containing potassium, sodium, and aluminum leads always to a residual liquid enriched in alkali-aluminum silicate. Bowen therefore referred to the $NaAlSiO_4$—$KAlSiO_4$—SiO_2 system as "petrogeny's residua system." In the equilibrium diagram (Figure 5.11) the $KAlSi_3O_8$—$NaAlSi_3O_8$ join divides the system into two portions. Compositions between this boundary and the SiO_2 apex approach granitic compositions; compositions below this boundary (i.e., toward the $NaAlSiO_4$—$KAlSiO_4$ join) approach the compositions of alkaline magmas which may crystallize as alkali feldspar plus nepheline or leucite. The most significant feature is the low-temperature trough within the 1100° isotherm. Fractional crystallization of any melt in this system will result in a residual liquid, the composition of which lies in this low-temperature region. Bowen pointed out that, if fractional crystallization has been of fundamental importance in the differentiation of magmas, then those igneous rocks that are the products of the crystallization of residual melts should have salic components with bulk compositions lying in this low-temperature trough. He was
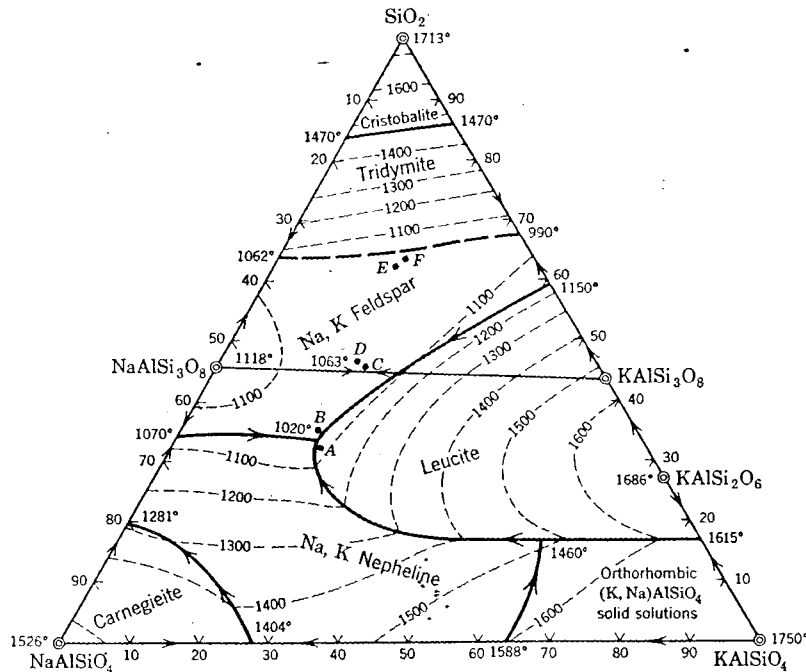
Figure 5.11 Equilibrium diagram for the system $NaAlSiO_4$—$KAlSiO_4$—SiO_2. The black dots represent plots of normative compositions (except anorthite) of the averages of 15 tinguaites (*A*), 25 phonolites (*B*), 32 alkaline syenites (*C*), 19 alkaline.trachytes (*D*), 546 granites (*E*), and 102 rhyolites (*F*). (Bowen, *Am. J. Sci.* **233**, 20, 1937, and Schairer, *J. Geol.* **58**, 514, 1950)

able to produce many such examples from among the rhyolites, trachytes, and phonolites, and plots of the salic portion of the average granite, rhyolite, syenite, etc:, all fall in the low-temperature region, as shown by Figure 5.11.

Further work on the $NaAlSi_3O_8$—$KAlSi_3O_8$—SiO_2 part of this system under water-vapor pressure has established these relations even more clearly (Figure 5.12). The phase relations in this system resemble those in the albite-anorthite-diopside system, there being two composition fields, one in which quartz is the first mineral to crystallize, the other in which alkali feldspar crystallizes first. These two fields are separated by the reaction curve *AB*. Under a water-vapor pressure of 1000 kg/cm^2, crystallization ultimately leads to a low-temperature trough around the middle part of *AB*. When the normative composition of analyzed granites is calculated in terms of SiO_2, $NaAlSi_3O_8$, and $KAlSi_3O_8$ and these points are plotted on a composition diagram, the frequency clearly reflects the position of this low-temperature trough; that is, the felsic portion of most granites corresponds in composition to the late liquid fraction produced by crystallization in the SiO_2—$NaAlSi_3O_8$—$KAlSi_3O_8$ system. This observation supports a process of fractional crystallization for the formation of
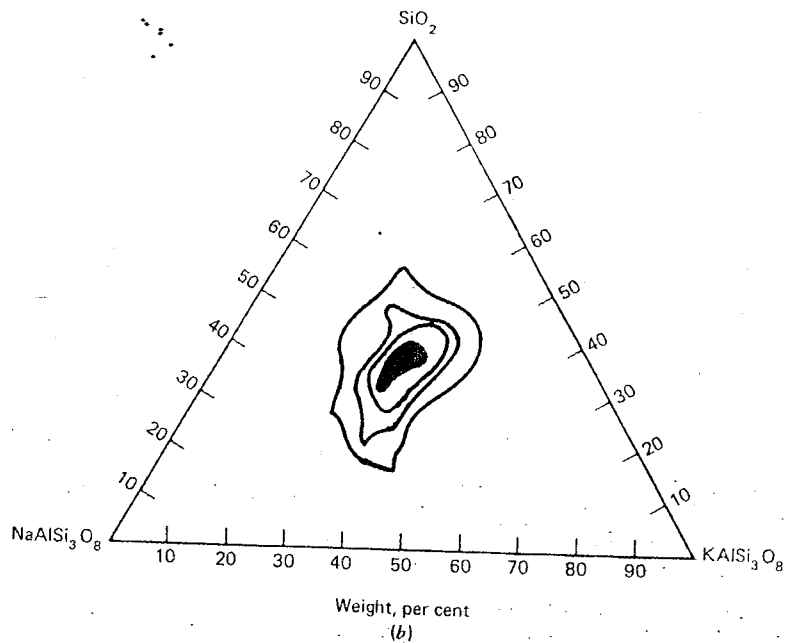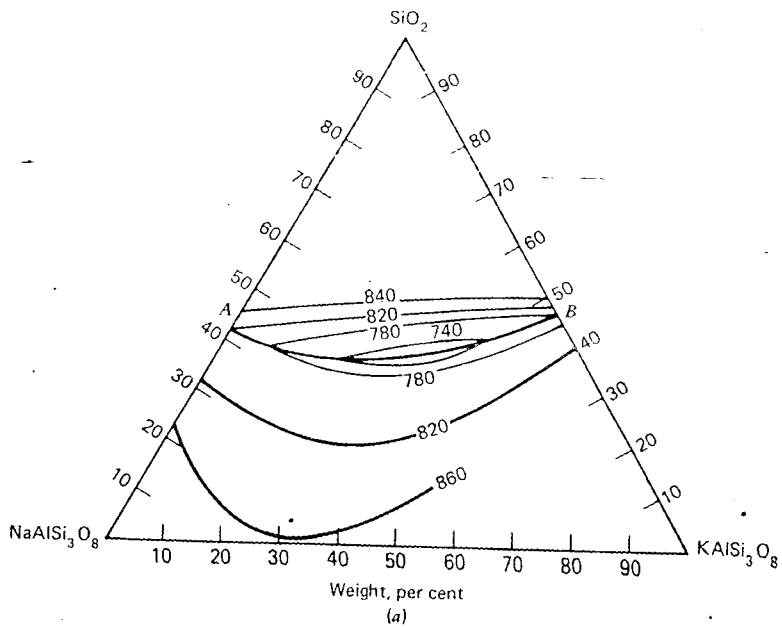
Figure 5.12 Composition of natural granites and its relation to the low-melting region of the system SiO_2—$NaAlSi_3O_8$—$KAlSi_3O_8$: (*a*) system under 1000 kg/cm² water vapor pressure (*AB* is the reaction curve separating the field of quartz from that of feldspar.); (*b*) contoured frequency-distribution diagram of the normative composition of over 500 analyzed granites containing 80 per cent or more of these components (over 90 per cent of these analyses fall within the solid black area). (Tuttle and Bowen, *Geol. Soc. Amer. Mem.* **74**, 1958)

most granites and is consistent with a magmatic origin for them. However, the evidence also shows that in fractional melting of rocks of appropriate composition the initial liquid would have a similar granitic composition.

THE CRYSTALLIZATION OF A MAGMA

From the study of crystallization in artificial silicate melts and the coordination of the results thereby obtained with the observations of igneous petrology, important conclusions have been established regarding magmatic crystallization. Simple eutectic crystallization, once believed to be common and important in magmas, probably never occurs. Nearly all rock-forming minerals are solid solution series. Crystallization of systems containing such compounds takes place over a range of temperature, and the phases separating from the melt have a considerable range of composition. The course of crystallization is dependent on the rate of solidification and the presence or removal of early-formed crystals, whereas in eutectic crystallization the rate of solidification has no influence, and the final condition is always the same, whether or not early-formed crystals are removed. Another significant feature in silicate systems is the frequent occurrence of incongruent melting, when one solid phase will be converted into a different phase by reaction with the liquid. Thus crystallization in a magma is characterized by reaction of two kinds: continuous reaction in a solid solution series, whereby early-formed crystals change uninterruptedly in composition by reaction with the melt, and discontinuous reaction, whereby an early-formed phase reacts with the melt to give a new phase with a different crystal structure and a different composition. This concept of *reaction* as the fundamental phenomenon of magmatic crystallization is due to Bowen and was developed by him into the *reaction principle*.

Bowen showed that the common minerals of igneous rocks can be arranged into two series, a discontinuous reaction series comprising the ferromagnesian minerals and an essentially continuous reaction series of the feldspars (Figure 5.13). In effect, each of the ferromagnesian minerals is itself a continuous reaction series, since all are solid solutions. The reaction series that Bowen set up, on the basis both of the laboratory study of silicate melts and the petrographic evidence, parallel the general sequence of magmatic crystallization as indicated by the petrology of the igneous rocks.

The petrological significance of the reaction principle may be illustrated by considering briefly the crystallization of a basaltic magma of such composition that olivine and bytownite are the first phases to form. As the temperature falls they react more or less completely with the melt and are converted into pyroxene and labradorite, and if no fractionation takes place the melt solidifies as a pyroxene-labradorite rock, a basalt or gabbro. If fractionation takes place and some of the early-formed olivine and bytownite is removed from the system, the reaction process will continue further, and the remaining melt will act upon pyroxene and labradorite to form hornblende and andesine. The greater the degree of fractionation, the more extensive the reaction process. With a high degree of fractionation, the whole reaction series is gone through, and the final liquid is a watery solution rich in silica.

Discontinuous reaction series				Continuous reaction series		
Minerals	Structure types	$\frac{Na, K}{Al}$	$\frac{Si}{Al}$	$\frac{Si}{Al}$	$\frac{Na, K}{Al}$	Minerals
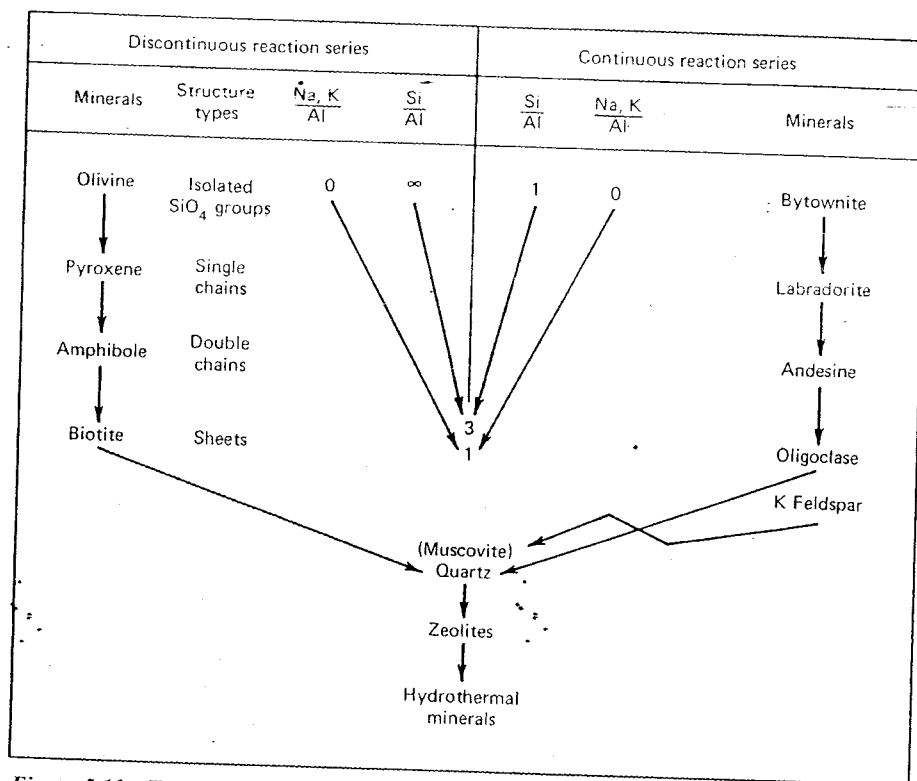
Figure 5.13 The reaction series.

At this stage it is of interest to examine some of the aspects of the continuous and discontinuous reaction series of Figure 5.13. First, the discontinuous reaction series is marked by increasing complexity of silicate linkage, the sequence being isolated tetrahedra–single chains–double chains–sheets. This sequence is also synonymous with greater size of the structural units, the unit cell volumes increasing as follows:

Forsterite	294 Å³
Diopside	434 Å³
Hornblende	925 Å³
Biotite	979 Å³

The latter part of the series is characterized by the introduction of fluorine and hydroxyl into the structures (hornblende and biotite), reflecting a higher concentration of volatiles as fractionation proceeds. Since biotite contains essential potassium, its crystallization in place of hornblende is probably conditioned in part by increasing concentration of potassium in the liquid.

This sequence of decreasing complexity of silicate linkage in passing up the discontinuous reaction series reflects an increasing thermal stability. Any of the linked structures lower in the series can be disintegrated into fragments of structures higher

in the series by heating. Thus with increasing temperature a mica sheet can conceivably be disintegrated into amphibole double chains, pyroxene single chains, and finally single tetrahedra, each step being accompanied by the production of a liquid residue. The melting of hornblende to give pyroxene plus liquid and of pyroxene to give olivine plus liquid has been observed in laboratory experiments.

Another noteworthy feature in the discontinuous reaction series is the increasing degree of replacement of Si by Al. In the olivine there is no evidence of any such replacement; in most magmatic pyroxenes the amount of replaced Si is generally low, but it is greater in the magmatic amphiboles, and in biotite at least one-fourth of the Si is always replaced by Al (the Si:Al ratio in igneous biotites ranges from 6:2 to 5:3). In contrast to this feature the Al:Si ratio in the continuous reaction series of the plagioclases shows a steady decrease from 1:1 in anorthite to 1:3 in albite and potash feldspar. The Na + K:Al ratio also shows a progressive change from 0 to 1 in both continuous and discontinuous branches.

The structural significance of the Al—Si substitution is evidently connected with the distribution of bond energies. If aluminum proxies for silicon in an SiO_4 group, the lower charge on the aluminum ion results in lesser neutralization of the negative charges on the oxygen anions, and this increment of negative charge leads to stronger bonding between the silicate units. Substitution of aluminum for silicon in a structure thereby tends to increase its disintegration temperature and consequently raises its position in the reaction series. This is true not only in the continuous reaction series, but also in individual minerals of the discontinuous reaction series. Aluminum substituting silicon in amphiboles and pyroxenes evidently increases their thermal stability.

The reaction series provides us with a concise statement of the segregation of the major elements during magmatic crystallization. The first-formed minerals, olivine and calcic plagioclase, are low in silica, and the liquid is thereby enriched in this component; the olivine is rich in magnesium, the plagioclase in calcium, and so the concentration of these elements in the liquid is decreased. The crystallization of olivine also changes the Mg:Fe ratio in the liquid, since this ratio is always higher in the crystals than in the liquid from which they separate. In effect, olivine and all the other ferromagnesian minerals are individual continuous series in which the early-formed crystals are magnesium-rich, the late-formed, iron-rich. As crystallization proceeds, pyroxene becomes the stable phase instead of olivine, and calcium may now be removed from the melt both as plagioclase and as augite. The liquid becomes relatively enriched in Na, K, and Si. As the concentration of calcium in the liquid falls the sodium content of the plagioclase progressively increases; with the appearance of hornblende as a stable phase in the discontinuous reaction series some sodium may be incorporated in this mineral also. Potassium remains in the liquid until a late stage, since it can be removed in appreciable amounts only in biotite and in potash feldspar.

According to the reaction principle, therefore, the fractional crystallization of a basaltic magma under suitable conditions can lead to the successive formation of more siliceous rocks until ultimately a granitic composition is reached. This sequence has been confirmed in many areas of igneous rocks, for example, the Oslo region, the

Caledonian plutonic rocks of western Scotland, and the batholith of southern California. Nevertheless, as Bowen himself pointed out, fractional crystallization is a very flexible process, and two magmas of approximately the same initial composition may produce very different rock types. One of the best documented examples of the fractional crystallization of a basaltic magma is the Skaergaard intrusion (Wager and Brown, 1967) In this intrusion differentiation has resulted in the late crystallization of iron-rich rock consisting essentially of plagioclase (An_{30}), olivine (Fa_{97}), and clinopyroxene close to hedenbergite. Here the trend of differentiation has been toward an enrichment in iron instead of an enrichment in silica. Osborn (1962) has shown that we can distinguish two distinct reaction series. The one leading to a silica-rich residual liquid, as described by Bowen, is characteristic of the igneous rocks of orogenic belts. The other reaction series leads to a residual liquid rich in FeO, and is typified by layered intrusions such as Skaergaard, the Bushveld intrusion, and the Stillwater complex. Each of these differentiation series is equally the product of fractional crystallization and reaction, but the very different results reflect a difference in the partial pressure of oxygen during crystallization. The original reaction series of Bowen describes what happens during the fractional crystallization of a basaltic magma under a relatively high oxygen fugacity or partial pressure. Under these circumstances much of the iron is removed as magnetite, the silica content of the liquid is thereby enhanced, and the residual liquid is silica-rich. On the other hand, under relatively low oxygen pressure little or no ferric iron is present to form magnetite and the FeO accumulates in the liquid until it is removed at a late stage by the crystallization of iron-rich olivine and pyroxene. It has been suggested that the oxygen is provided and controlled by the continuous input of water and possibly CO_2 from surrounding rocks. Either H_2 or CO would be lost from the open system to account for the chemical balance. This theory is supported by oxygen isotope data, which indicate a continuous exchange of oxygen with a reservoir of water in the earth's crust.

Recent work on the crystallization of natural and synthetic melts under high pressures has greatly extended our understanding of the scope and variety of magmatic differentiation. For example, the incongruent melting of $MgSiO_3$ to give Mg_2SiO_4 and a silica-rich liquid is eliminated at quite moderate pressures, indicating that although a reaction relation between olivine, orthopyroxene, and liquid may be an important factor in near-surface magmatic differentiation, it will be greatly modified in the deeper levels of the crust. Green and Ringwood (Table 5.1) have demonstrated experimentally that the field of primary crystallization of orthopyroxene from basaltic magma is substantially increased at high pressure. With increasing pressure the amount of olivine decreases and that of orthopyroxene increases. The composition of the orthopyroxene changes with increasing pressure, becoming progressively richer in aluminum. Green and Ringwood have also studied the crystallization of a basaltic glass at 1100° and different pressures. Up to 15 kb the glass crystallizes to pyroxene and plagioclase, but the plagioclase decreases in amount and becomes more sodic with increasing pressure, evidently through the incorporation in the pyroxene of calcium and aluminum from the anorthite component. Above 15 kb garnet (pyrope-almandine) appears and increases in amount with increasing pressure, being formed

Table 5.1 Crystallization of Basaltic Melt under Differing
Conditions of Temperature and Pressure

P (kb)	T	Phases Present	Remarks
10	1250	ol + opx + liq	ol $\gg$ opx
12½	1250	ol + opx + ?cpx + liq	opx > ol
15	1300	ol + opx + cpx + liq	opx $\gg$ cpx > ol
20	1300	opx + liq	

ol = olivine,
opx = orthopyroxene,
cpx = clinopyroxene.
Green and Ringwood, *Nature* **201**, 1276, 1964.

by the breakdown of aluminous pyroxene. Plagioclase disappears above 20 kb, and quartz appears, evidently produced by the decomposition of the albite component into jadeite going into solid solution in the pyroxene. This sequence is petrologically equivalent to the sequence gabbro—garnet granulite (garnet, pyroxene, sodic plagioclase)—eclogite. The different behavior under high pressure of the anorthite and albite components of plagioclase is particularly noteworthy. In this connection potash feldspar provides a further contrast; there is no corresponding compound to jadeite ($KAlSi_2O_6$—leucite—is a feldspathoid and unstable above about 2 kb), and potassium cannot be accommodated in the pyroxene structure. Under high pressures in the lower crust and upper mantle potassium is probably contained in hornblende or phlogopite.

Thus the solidification of a magma, although governed by the simple principles of phase equilibria and fractional crystallization, is capable of a remarkable versatility under the variety of conditions in the crust and upper mantle. The distribution of the major elements is controlled by the sequence of mineral separation as the temperature drops and as the composition of the melt changes by the removal of some elements in the solid phases. The principles governing magmatic crystallization are, however, essentially descriptive, not explanatory. They describe how the major elements distribute themselves among the different minerals but not why they act in this way, nor do they tell us anything regarding the fate of the minor elements. The explanation must ultimately depend on thermodynamic considerations, the energy changes involved in assembling ions of differing size and charge in particular crystal lattices.

THE THERMODYNAMICS OF MAGMATIC CRYSTALLIZATION

The thermodynamics of magmatic crystallization is controlled by the nature of the ions present, their concentration, the temperature and pressure, and the type of crystal lattices formed. In effect, the question posed is, "How do changes in the composition and physical conditions of the magma affect the solubility of the many different possible compounds?" According to thermodynamics, a solid A will crystallize from a liquid containing A if the chemical potential of A is less in the solid than in

the liquid. However, although we can state the problem, we cannot solve it, since the necessary thermodynamic data are not available.

Crystallization is essentially an ordering phenomenon. In a crystal the degree of order is great and the ions are arranged in a regular lattice, whereas in a melt the arrangement of the ions, although not entirely random, is of a lower degree of order. From the viewpoint of thermodynamics magmatic crystallization is characterized by the energy and entropy changes involved in removing ions from the melt and packing them in an orderly fashion in a crystal lattice. The fate of an element during magmatic crystallization is linked with its concentration in the magma and the nature of the structural lattices that may form. The silicon and aluminum content of the magma and the temperature are the factors controlling the sequence of crystal lattices. These crystal lattices act as a sorting mechanism for the cations. A cation can enter a crystal lattice only if it is of suitable size and can attain its appropriate coordination number. Since in general a number of different ions can satisfy this requirement, that ion enters the lattice in largest amount, relative to its concentration in the liquid, which holds its position in the lattice with greatest tenacity. From studies of crystal structures and independently of energy considerations, Goldschmidt formulated the following empirical rules as a general guide to the course of an element during liquid → crystal formation in a multicomponent system:

1. If two ions have the same radius and the same charge, they will enter a given crystal lattice with equal facility.
2. If two ions have similar radii and the same charge, the smaller ion will enter a given crystal lattice more readily.
3. If two ions have similar radii and different charge, the ion with the higher charge will enter a given crystal lattice more readily.

The rules essentially state that the cation that forms the strongest chemical bond will preferentially fill a lattice site.

Goldschmidt's rules have wide application in the geochemistry of igneous rocks; we have already seen examples in considering the reaction principle. Calcium enters the feldspar lattice more readily than does sodium on account of its higher charge and so is concentrated in the early-formed plagioclase. The magnesium ion is somewhat smaller than the ferrous ion, and magnesium is always concentrated in the early-formed ferromagnesian minerals, whether olivine, pyroxene, amphibole, or biotite. However, Goldschmidt's rules have had their greatest utility in predicting the order of removal from a magma not only of the major elements but of the minor elements also.

Chemical variations reflecting the sequence of crystallization of minerals from a magma or among selected rocks from different areas may be illustrated by means of variation diagrams. Variation diagrams are graphs showing chemical constituents of interest plotted against a selected variable. The variable is most often selected to reflect the evolutionary trend of a magmatic rock series. Commonly the SiO_2 content of individual rocks is used, since SiO_2 tends to increase with mineralogical fractionation. For general considerations it is also possible to use common rock names placed on the graphs in the position of average SiO_2 content. Such a diagram is shown in
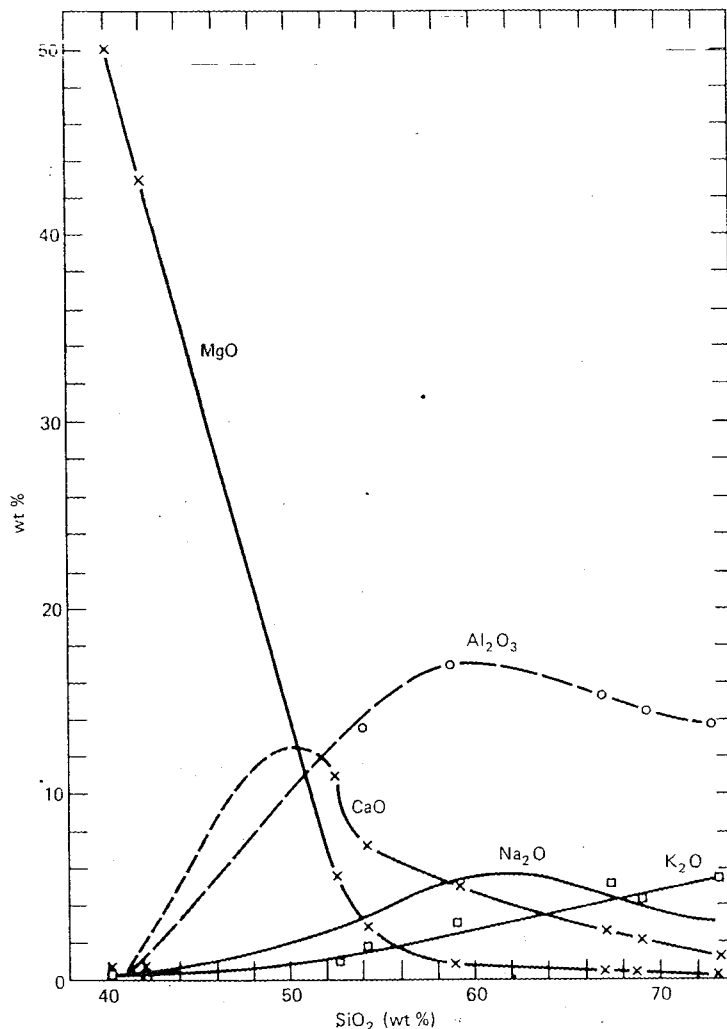
Figure 5.14 Variation diagram for the USGS standard rocks.

Figure 5.14 illustrating the variations of some of the specific elements discussed later in this chapter. Other variables have been used to reflect evolution of a rock series. A system often used is the Larsen diagram where the variable ($\frac{1}{3}$ SiO_2 + K_2O) — (CaO + MgO + FeO) is used. Triangular variation diagrams may also be used to plot the mutual relations of three components or sets of components recalculated to 100%. Commonly used is the system (Na_2O + K_2O), (FeO + Fe_2O_3), and MgO. Rocks derived from the same initial magma generally show smooth variation diagrams for individual components. Exotic rocks may plot away from the smooth curves, indicating their different origin.

MINOR ELEMENTS IN MAGMATIC CRYSTALLIZATION

When a minor element has the same charge and an ionic radius similar to a major element, we speak of it as being *camouflaged* in the crystal lattice containing the major element. Thus Ga^{3+} (0.62 Å) is camouflaged in aluminum minerals, and Hf^{4+} (0.78 Å) is camouflaged in zirconium minerals. When a minor element has a similar ionic radius but a higher charge than that of a major element (or the same charge but a lesser radius), it is said to be *captured* by the crystal lattice containing the major element. Thus Ba^{2+} (1.36Å) is captured by potassium minerals (K^+ 1.38Å). Finally, when a minor element has a similar ionic radius but a lower charge than that of a major element (or the same charge but a greater radius), it is said to be *admitted* into the crystal lattice containing the major element. Thus Li^+ (0.74Å) is admitted into magnesium minerals. In capture and admission of ions of different charge the charge balance is maintained by concomitant substitution elsewhere in the crystal lattice.

Goldschmidt's rules are a useful guide to the distribution of the elements during magmatic crystallization, but they are not universally valid and have been criticized on that account. The major source of this lack of universal validity seems to lie in the fact that the bonding in most minerals is not exclusively ionic, whereas the rules are predicated on a purely ionic basis. This has been carefully considered by Ringwood (1955), who shows that the electronegativity of an element, a measure of its tendency to form covalent bonds, has an important influence on the extent to which it will proxy for another element of similar size. Elements of the first B subgroup of the periodic table have considerably higher electronegativities than corresponding elements of the A subgroups, and this must be considered when comparing diadochy between elements belonging to different subgroups (Table 5.2). Ringwood expresses the electronegativity factor in the form of a rule: Whenever diadochy in a crystal is possible between two elements possessing appreciably different electronegativities the element with the lower electronegativity will be preferentially incorporated because it forms a stronger and more ionic bond than the other. In practice this rule is found to apply to substitutions involving elements differing in electronegativity by more

Table 5.2 Electronegativities of the Commoner Cations

Cation	Electronegativity	Cation	Electronegativity
Rb^+	0.8	Mn^{2+}	1.4
K^+	0.8	Cr^{3+}	1.6
Ba^{2+}	0.85		
Na^+	0.9	Fe^{2+}	1.65
Sr^{2+}	1.0	Ni^{2+}	1.7
Ca^{2+}	1.0	Co^{2+}	1.7
Li^+	1.0	Zn^{2+}	1.7
Rare earths	1.05–1.2	Fe^{3+}	1.8
Mg^{2+}	1.2	Cu^+	1.8
Sc^{3+}	1.3	Cu^{2+}	2.0
V^{3+}	1.35		

than 0.1. The significance of bond type on the distribution of the elements has been extensively reviewed by Ahrens (1964).

In order to evaluate the movement of an element between a melt and a crystalline phase, consideration must be given to its chemical nature in the liquid as well as in the solid state. Elements that readily form complexes with oxygen or other anions often will not enter crystal lattices until the activity of their complex form is high enough to allow new mineral phases to crystallize. Cations with ionic charges of four or larger and some small-radius trivalent cations occur as complexes with oxygen or hydroxyl anions in magmas. Included in this group is silicon in the form of SiO_4^{4-}, the fundamental building block of silicate minerals. A simple measure of the probability of an element to form a complex is its ionic potential or ratio of ionic charge to ionic radius. Elements with high ionic potential values tend to form complexes. This parameter is also useful in determining the species of dissolved elements in water solutions and is described in more detail in Chapter 6. In common magmas elements with ionic potentials between 0.6 and 2.5 may be considered to occur as free ions occupying holes in the network of silica polymers. Elements with ionic potentials between 2.5 and 4.8 maintain an equilibrium between complex and noncomplex species. The concentration of each species in equilibrium is determined by the composition of the magma. In most magmas they behave as free ions but in oxygen-rich solutions they form complexes. Elements with ionic potentials greater than 4.8 form complexes that may or may not be linked to the silica network. Elements forming nontetrahedral complexes are not readily incorporated into silicate minerals and hence will become concentrated in residual magmas. Examples of such complexes include carbonate CO_3^{-2} and borate BO_3^{-3} species. If elements form tetrahedral complexes, the larger the charge on the central cation the less readily is a complex accepted into a silicate structure. Hence, complexes in which the central charge is greater than four become concentrated in the fluid phase of magmas. Such complexes as phosphate PO_4^{-3}, tungstate WO_4^{-2}, and sulfate SO_4^{-2} fall into this category. Likewise the larger the size of the central ion in a tetrahedral complex the less readily it is accepted into a silicate structure. Thus GaO_4^{-4} and GaO_4^{-5} tend to accumulate in residual fluids relative to SiO_4^{-4} and AlO_4^{-5}. Minerals containing these rejected complexes are commonly found in pegmatites and hydrothermal deposits.

A more refined approach to the behavior of the first set of transition elements of B subgroup elements is provided by crystal field theory. These transition metals possess an incomplete inner d-electron shell. When a transition metal is surrounded and influenced by the electrical field of anions, the ion's spherical bonding symmetry and energy change depending on the coordination of the anions around it. A detailed discussion of the extra or crystal field stabilization energy and its application to geochemistry is found in Burns (1970). The crystal field stabilization energy varies with transition metals depending on whether they are in the divalent or trivalent states and whether they are in octahedral or tetrahedral coordination. As a general rule, the ionic sizes of the transition metals require that they have octahedral coordination in silicate crystals and tetrahedral coordination in the parent magma. An octahedral site preference energy can be derived that is a partial measure of the preference of

a transition metal for the crystalline rather than the liquid state. The order of octahedral site preference energy for divalent and trivalent ions is as follows:

$$M^{++} \text{ ions } Ni > Cu > Co > Fe > Mn = Zn = Mg$$
$$M^{+++} \text{ ions } Cr > Co > V > Ti > Fe = Sc = Ga$$
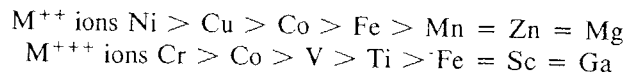
These sequences match the observed order of depletion of trace elements in the Skaergaard Intrusion illustrated by a variation diagram in Figure 5.15. Figure 5.15 shows the success of the above predictions with the notable exception of copper. It has been suggested that copper does not enter easily into the silicate minerals because of a predictable but unusual lattice distortion, energy destabilization effect.

The fate of the individual elements during magmatic crystallization is now discussed briefly, and is illustrated by Tables 5.3 and 5.4.
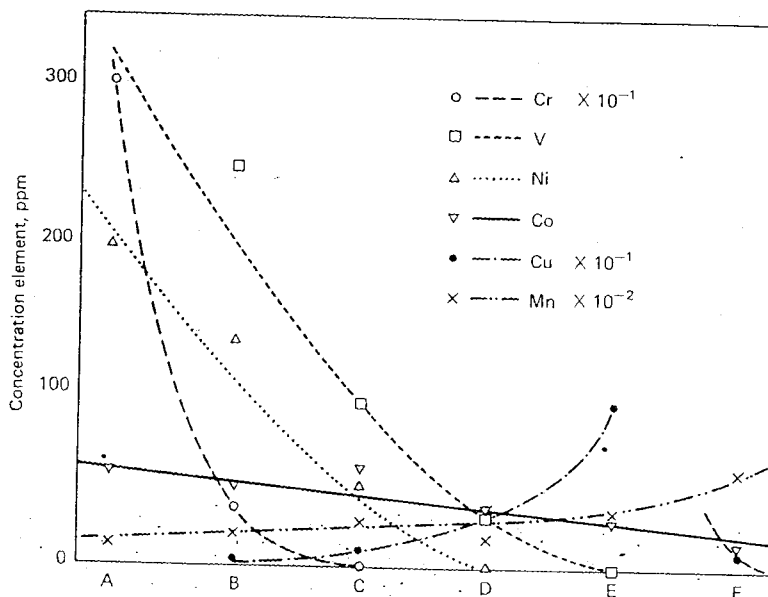
Figure 5.15 Distribution of trace elements in Skaergaard Complex pyroxenes from selected rock types:
A. The earliest olivine gabbro exposed
B. Hypersthene olivine gabbros
C. Olivine-free gabbros
D. Hortonolite ferrogabbros.
E. Ferrohortonolite gabbros
F. Fayalite ferrogabbros
(From Curtis, *Geochim. et. Cosmochim. Acta* vol. 28, p. 398, 1964).

Table 5.3 Element Distribution in Minerals of Granodiorite Gneiss, West Greenland*

Element	Plagioclase	Microcline	Biotite	Hornblende	Titanite	Apatite
Cs	0.2	0.4	13	<0.1	<0.1	<0.1
Rb	30	150	120	7.7	<0.3	<0.3
Ba	250	3700	1400	28	2.0	3.2
Pb	16	20	3.3	7.3	4.1	4.9
Sr	280	250	4.0	13	4.8	120
La	0.9	0.5	1.1	15	58	44
Y	0.5	0.4	0.4	50	500	250
Nb	<0.2	<0.2	5.6	11	80	1.0
Mn	—	—	2000	3300	1200	—
Ti	—	—	1.55%	6900	21.8%	—

After Mason, *Geol. Surv. Greenland,* Report No. 71, 1975.
*Figures are ppm unless otherwise stated.

Table 5.4   Element Distribution in Minerals of Gabbro, Muzzle River, New Zealand*

Element	Plagioclase	Orthoclase	Hornblende	Biotite	Apatite	Titanite
Rb	42	>200	5	>200	—	—
Ba	510	810	440	1600	6	150
Pb	2	3	<1	60	<1	260
Sr	800	1200	600	420	700	190
La	19	7.8	22	21	1100	400
Y	5.8	1.5	35	11	210	160
Sc	14	8	32	27	<1	>100
Zr	45	60	330	160	16	7000
Nb	3	2	65	60	1	2000
Sn	<1	<1	2	8	1	23
Mn	40	70	1200	2100	160	700
Ti	200	200	3.2%	3.4%	100	23.8%
V	8	8	120	100	20	110
Ga	14	17	20	60	—	—

After Mason, *New Zealand J. Geol. Geophys.* 15, 465, 1972.
*Figures are ppm unless otherwise stated.

Cesium

This is the largest cation, and potassium is the only common cation it can readily replace. Among the rock-forming minerals, it is concentrated in biotite relative to potash feldspar (Table 5.3), evidently because the K site in biotite is considerably larger than that in feldspar. However, during magmatic crystallization most of this element evidently remains in the liquid until a late stage of crystallization, and it is found concentrated in granite pegmatites in micas and potash feldspar. Maximum contents of Cs_2O of 1.9% in lepidolite and 0.7% in microcline have been recorded from the Varuträsk pegmatite in Sweden. Normally the concentration of Cs does not

reach high enough values to form separate minerals, but high concentrations of Cs in some pegmatite liquids may lead to the formation of pollucite, $CsAlSi_2O_6$.

Rubidium

The rules of trace element distribution predict that Rb should replace K and this is found to be true. Rubidium forms no minerals of its own, being always incorporated in potassium minerals; in igneous rocks it is in muscovite, biotite, and potash feldspar (Tables 5.3 and 5.4). Since Rb^+ is considerably larger than K^+, rubidium is admitted into potassium minerals, and accordingly the Rb:K ratio increases with increasing differentiation; this ratio is highest in pegmatite feldspars and micas.

Barium

The only major element of comparable ionic size is potassium, and barium appears, therefore, in biotite and potash feldspar. Because of its higher charge barium should be captured by potassium compounds. The available data indicate that this is generally true, barium being relatively enriched in early-formed potassium minerals. Barium may be accepted in the plagioclase structure, the maximum recorded being about 0.1%, and to some extent in hornblende, which has a structure site large enough to accommodate it.

Lead

Traces of lead, the most abundant of the heavy elements, are found in the silicate material of many igneous rocks, particularly in granites. It is evidently present as the Pb ion (1.18 Å), diadochic with K^+. From the ionic charge alone it might be expected that the lead in a magma would tend to be captured by potassium minerals, but the much greater electronegativity of lead evidently results in a weakening of this tendency, and lead is admitted rather than captured by potassium minerals.

Strontium

The size of the strontium ion indicates that it can proxy for either calcium or potassium, being admitted to calcium minerals (larger radius) or captured by potassium minerals (higher charge). The data indicate that strontium in igneous rocks is present mostly in the plagioclase and potash feldspar and that its concentration in specific minerals increases as crystallization proceeds. Its abundance in igneous rocks is somewhat variable, generally being higher in basalts and gabbros than granites. Admittance in place of calcium is the dominant process of removal of strontium from the magma. However, this process does not operate with all calcium minerals; strontium is present in only insignificant amounts in augite, suggesting that the pyroxene structure does not readily accommodate the strontium ion. Strontium has similiar overall geochemical characteristics to barium.

The Rare Earths (including Y)

The charge and the comparatively large radius (Lu 0.85 A—La 1.06Å) of the rare earths, coupled with their general low concentration, suggest that they would show relatively little tendency to replace the major elements during magmatic crystallization, and this is borne out by the concentration of rare earths as individual minerals in pegmatites. However, some replacement of Ca^{2+} by rare earths can take place in apatite, and this mineral is the principal carrier of rare earths in most igneous rocks. Titanite, $CaTiSiO_5$, when present, is also a significant carrier of these elements (Tables 5.3 and 5.4). In granites and pegmatites formed at sufficiently low temperatures for the epidote structure to be stable we find the mineral allanite, in which some of the calcium of epidote is replaced by rare earths.
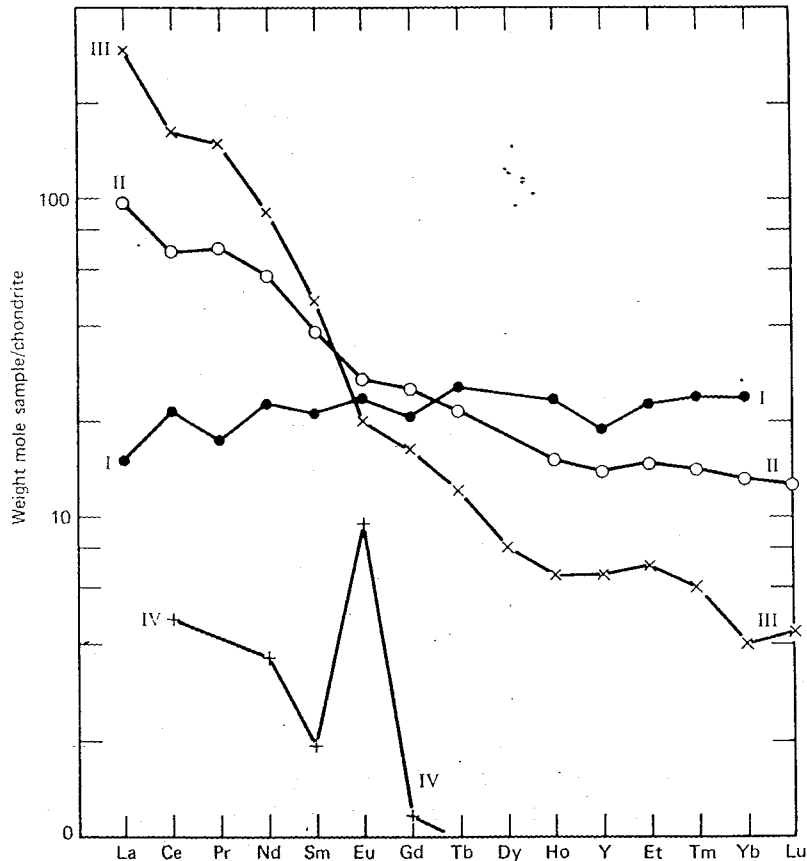
Figure 5.16 Chondrite-normalized rare earth concentrations in (I) oceanic basalt from the Mid-Atlantic ridge, (II) a composite of continental basalts, (III) granite G-1, (IV) anorthosite, Quebec. Note that Y is plotted between Ho and Er, since its ionic radius is essentially identical with these elements.

In view of the extreme difficulty of separating the individual rare earth elements in the laboratory. it was long assumed that marked fractionation of these elements was unlikely to occur during magmatic differentiation; in any event, the problem could hardly be investigated until satisfactory methods for the analysis of these elements at trace levels were developed. Such methods have become available, and the wealth of new data has shown that the rare earths are indeed fractionated during magmatic processes, and in ways that throw considerable light on these processes. Because of their smooth change in ionic radii the rare earths are sensitive indicators to different igneous processes. A few illustrations are provided in Figure 5.16. In this diagram the data are chondrite-normalized; that is, the concentrations for the individual elements are divided by the corresponding chondritic abundances, considered to be a close approximation to cosmic abundances. This procedure in effect brings out the degree of fractionation from a primordial distribution pattern and smoothes out the odd-even zig-zag pattern of absolute rare-earth element abundances. The least-fractionated pattern in Figure 5.16 is that for an oceanic basalt dredged from the Mid-Atlantic ridge, in which the individual elements are uniformly enriched at 15 to 25 times chondritic abundances. Continental basalts show a somewhat fractionated pattern, with relative depletion with increasing atomic number. In general, the degree of this fractionation increases with increasing magmatic differentiation, granites such as G-1-showing the greatest effect. Presumably this relative depletion in the heavier rare earths must be matched by relative enrichment of these elements in some complementary material. Some peridotites show a relative enrichment in the heavier rare earths. Clinopyroxene and especially garnet are minerals that selectively concentrate the heavier rare earths, and the retention of these minerals in the upper mantle may explain the fractionation of these elements seen in igneous rocks.

The data for the anorthosite are included in Figure 5-16 to illustrate the unique behavior sometimes shown by europium. Anorthosite is a rock consisting essentially of plagioclase feldspar, and the Eu anomaly is clearly linked with the feldspar. Under moderate reducing conditions Eu is unique among the rare earths in existing stably in the divalent state; in this state its ionic radius is close to that of Sr, and it is then concentrated in Sr-bearing minerals, which in igneous rocks are the feldspars. Thus feldspar-rich rocks commonly show a positive Eu anomaly, and this anomaly is even more marked in feldspar-rich meteorites and lunar rocks, which crystallized under more reducing conditions than terrestrial igneous rocks.

Manganese

Manganese is present in magmas as the Mn^{2+} ion and in this form may be expected to proxy for Fe^{2+} or for Ca^{2+}. However, manganese is much more electronegative than calcium and probably on this account is seldom found replacing this element (except in the apatite of pegmatites). The manganese in igneous rocks replaces ferrous iron, and a relative increase in the Mn:Fe ratio has been noted in later differentiates, indicating that the larger size of the manganese ion causes it to be admitted to ferromagnesian minerals. There is no crystal field stabilization energy for manganese. Table 5.3 shows clearly the absence of manganese from the calcium-bearing

plagioclase and its essentially uniform distribution in the ferromagnesian minerals hypersthene and olivine.

Zirconium

The combination of high charge and comparatively high radius (0.79 Å) sets zirconium apart from any of the major elements of igneous rocks. On this account zirconium does not enter into the common rock-forming minerals to any degree but appears in a specific phase, usually zircon. Zircon is most abundant in later differentiates, evidently because the original zirconium concentration of a magma is generally less than saturation for zircon.

Hafnium

Hafnium, having the same charge as Zr and about the same radius, always occurs camouflaged in zirconium minerals, and the Zr:Hf ratio remains almost constant throughout any process of fractional crystallization, at about 50. However extreme differentiation can result in an enrichment of Hf relative to Zr.

Scandium

Scandium has a radius close to that of ferrous iron, and, in view of its higher charge, scandium would be expected to be captured by ferromagnesian minerals. This is true for magmatic pyroxenes, which generally show a relative concentration of scandium (Table 5.3); it may also be present in hornblende and biotite, and is notably enriched in titanite (Table 5.4). Scandium is not concentrated in the even earlier-formed olivines, evidently because of the difficulty in balancing the excess positive charge thus introduced by other suitable replacements.

Cobalt

The bivalent cobalt ion is practically the same size (0.74Å) as the ferrous ion (0.77Å), and cobalt should thus be camouflaged in ferrous compounds. It is found, however, that the Co:Fe ratio is greatest in early-formed minerals and decreases steadily with increasing fractionation. Due to crystal field stabilization, the effective radius of cobalt is, therefore, somewhat less than the radius given above and is apparently almost identical with that of Mg, since Nockolds and Allen found the Co:Mg ratio to be almost constant throughout a rock series. The major part of the cobalt in a magma is removed in the early-formed magnesian minerals, especially olivine.

Nickel

The nickel ion has essentially the same radius and the same charge as magnesium, and therefore should be camouflaged in magnesium minerals. However, the ratio Ni:Mg is highest in early-formed crystals (especially olivine) and shows a steady

decline in the later-formed rocks and minerals. Nickel has the strongest crystal field stabilization energy of the common divalent ions and thus competes successfully with Fe^{2+} and Mg^{2+} for octahedral lattice sites.

Lithium

On the basis of chemical properties it might be expected that lithium would follow the other alkali elements in magmatic crystallization. However, ionic size, not chemical properties, is decisive in crystallization, and the lithium ion is very much smaller than any of the other alkali ions (Li$^+$0.74 Å, Na$^+$ 1.02 Å, K$^+$ 1.38Å). Hence lithium follows magnesium, since the ionic sizes are nearly identical; because the lithium ion has a lower charge than the magnesium ion it should be admitted into magnesium minerals, and this is found to be true. The Li:Mg ratio shows a steady increase in later-formed rocks and minerals. Strock, who made careful measurements of this ratio for different igneous rocks, suggested that it could be used as an index to the stage of differentiation reached by a given rock. Lithium is found in the pyroxenes, amphiboles, and particularly the micas. However, a considerable amount remains in the liquid until a very late stage of differentiation, since pegmatites often show a particular concentration of lithium, which in the praetical absence of magnesium forms individual minerals, such as lepidolite, spodumene, amblygonite, and petalite.

Vanadium

Vanadium is probably present in magmas as the V^{3+} ion. It is largely removed in magnetite, in which it proxies for Fe^{3+}; its ionic radius is greater than ferric iron, but its electronegativity is much less and its crystal field stabilization energy higher. These later two factors evidently being responsible for the enrichment of vanadium in early-formed magnetite. Vanadium also occurs in pyroxenes, amphiboles, and biotite and has been noted in marked concentration in aegirine, a mineral with a high Fe^{3+} content.

Chromium

Chromium is also present in magmas as the Cr^{3+} ion. The radius of this ion is very close to that of Fe^{3+}, but chromium shows a high degree of preferential concentration relative to ferric iron, being largely removed from a magma in the early stage of crystallization as chromite. This can be ascribed to the strong crystal field stabilization energy of Cr^{3+} relative to that of Fe^{3+}. Chromium is also enriched in pyroxenes, especially those from ultrabasic rocks.

Titanium

In igneous rocks titanium is present mainly as ilmenite. It can replace Al in six-coordination and appears therefore in pyroxene, hornblende, and biotite; it is probably captured by such minerals on account of its higher charge (Ti^{4+}—Al^{3+}). Tita-

nium does not appear in muscovite, probably because in highly siliceous magmas it is removed as titanite.

Gallium

Gallium, with the same ionic charge and radius close to that of aluminum, is camouflaged in aluminum-bearing minerals. The somewhat larger size of the gallium ion (Ga^{3+} 0.62 Å, Al^{3+} 0.51 Å) suggests that gallium would tend to be more abundant in later-formed aluminum minerals. Measurements of the Ga:Al ratio in igneous rocks and their minerals show that this ratio is nearly constant, indicating effective camouflage of gallium in aluminum minerals. There is a tendency for the Ga:Al ratio to increase in the later differentiates, and Ga is enriched in pegmatite feldspars and micas.

Germanium

The germanium ion has the same charge and a somewhat higher radius than silicon (Ge^{4+} 0.53 Å, Si^{4+} 0.42 Å). Germanium replaces silicon, and measurements of the Ge:Si ratio in silicates generally show little variation, indicating that it is effectively camouflaged in such minerals; there is, however, evidence of some degree of concentration of germanium in late differentiates.

The remaining lithophile elements are those that do not substitute the major elements because of a great difference in ionic radius and ionic charge. On account of low concentration in the original magma they remain in solution and are enriched in the residual liquid of magmatic crystallization. These elements are B^{3+} (0.12 Å), Be^{2+} (0.27 Å), W^{6+} (0.60 Å), Nb^{5+} (0.64 Å), Ta^{5+} (0.64 Å), Sn^{4+} (0.69 Å), Th^{4+} (1.04 Å), U^{4+} (1.00 Å), and Cs^+ (1.70 Å) and also the rare earths (0.85 − 1.06 Å), Li^+ (0.74 Å), and Rb^+ (1.49 Å). They are concentrated in pegmatites, the only economic source of many of them.

During the final crystallization of the residual liquid these elements may enter the structure of a common mineral (Rb in microcline) or a rarer mineral (Mn in apatite); they may enter or be trapped by a much less appropriate mineral structure (Sn in muscovite), followed in some instances by transfer to an exsolved mineral; they may concentrate in the residual fluid till a specific mineral is formed (Cs in pollucite, Be in beryl); or they may be bound to the surface of some mineral or minerals by adsorption. This last process evidently is effective in the crystallization of granitic rocks and results in the attachment of the ions of minor and trace elements to the crystal surfaces of common silicate minerals. The binding by adsorption is relatively weak, as demonstrated by experiments in which considerable amounts of minor and trace elements have been removed from granites by leaching with dilute acids.

It might be expected that the chalcophile elements in a magma would associate with sulfur and lesser amounts of arsenic, antimony, bismuth, selenium, and tellurium to form sulfides and related compounds. However, the only sulfides commonly present in normal igneous rocks are pyrite and pyrrhotite. The sulfur content of mag-

Table 5.5 The Content of Minor Elements in Differentiates of the Skaergaard Intrusion Compared to That in the Initial Magma*

	Initial Magma	Gabbro Picrite and Eucrite (earliest)	Olivine Gabbro	Olivine-Free Gabbro	Ferrogabbros		Hedenbergite Granophyre	Granophyre (latest)
Rb	—	—	—	—	—	—	—	—
Ba	40	25	25	45	—	—	30	200
Sr	300	200	700	450	50	150	450	1700
La	—	—	—	—	700	450	500	300
Y	—	—	—	—	—	—	25	150
Zr	40	40	35	25	20	125	175	200
Sc	15	7	20	15	10	100	500	700
Cu	130	70	80	175	400	—	—	—
Co	50	80	55	40	40	200	500	20
Ni	200	600	135	40	—	5	10	4
Li	3	3	2	3	—	—	5	5
V	150	170	225	400	3	15	25	12
Cr	300	700	175	—	15	—	—	12
Ga	15	12	23	15	20	20	40	30
Proportion of rocks as percentage		65	14	10	7½	2½	½	½

From Wager, *Observatory* 67, 103, 1947.
*Figures in ppm.

mas is relatively low, and nearly all combines with the abundant iron. Many of the chalcophile elements become predominantly incorporated in silicates rather than in sulfides. This can be partly accounted for by the low thermal stability of most sulfides, so that sulfides, except for pyrite and pyrrhotite, characteristically segregate at a late stage and are ultimately deposited in veins formed at comparatively low temperatures. The relative concentration of sulfur and chalcophile elements probably increases in the late liquid fraction of a crystallizing magma as a result of the removal of the lithophile elements in the silicate minerals.

The work of Wager and Mitchell (1951) on the Skaergaard intrusion has given a useful account of the distribution of minor elements in different rocks which are clearly related on the basis both of field exposures and of their petrology. The sequence of differentiation is from gabbro to granophyre, and the amounts of the minor elements in various rocks of this sequence are given in Table 5.5. The amounts assigned to the initial magma are based on the analysis of material from the chilled margin of the intrusion. The correlation between the results presented in Table 5.5 and predictions on the basis of Goldschmidt's rules is striking and provides independent support for the belief that the rocks of the Skaergaard intrusion are the result of differentiation by fractional crystallization.

FRACTIONAL CRYSTALLIZATION AND MELTING

Experimental and theoretical considerations of the behavior of trace elements in fractional crystallization and partial melting add valuable information on the origin and chemical evolution of magmas. Models of trace element behavior are based on the selection of solid-liquid partition coefficients, K, of trace elements. When a mass dm of a liquid magma crystallizes, a portion dx of the trace element enters into the solid. Under equilibrium conditions

$$\frac{dx}{dm} = K\frac{x}{m}$$

where x is the amount of trace element remaining in the magma of mass m. Integration yields an expression for the concentration of the trace element C_m in the magma

$$C_m = C_0\,F^{K-1}$$

where C_0 is the trace element concentration in the original magma and F is the fraction of the magma not yet crystallized. The concentration of the element in the crystallizing solid is

$$C_s = K\,C_{\text{liquid}}F^{K-1}$$

For a homogeneous solid in which the most recently crystallized fraction is distributed throughout earlier formed crystalline material or a bulk solid is taken for chemical analysis, a mean value of C_s can be obtained from

$$\overline{C_s} = C_{\text{liquid}}\,\frac{1-F^K}{1-F}$$

Under conditions of constant K curves of the concentration of an element in the solid phase may be determined as shown in Figure 5.17. Note that with high values of K a trace element will be rapidly depleted in the magma while for low values of K a trace element will be greatly enriched in the residual liquid phase. In real systems such consideration are complicated by the fact that K may change with temperature and compositions of the solid and magma and the successive crystallization of new phases will require simultaneous consideration of two or more partition systems. If partial remelting of a solid phase takes place in the evolution of rock-magma system, additional variables must be considered.

Of equal interest to the crystallization of magmas is the formation of magmas by the partial melting of rocks. The first liquid formed has a concentration of a trace element given by

$$C_1 = \frac{1}{K}\,C_{0\text{solid}}\,(1-F)^{(1/K-1)}$$

where K is again the solid-liquid partition coefficient. K is usually replaced by a bulk distribution coefficient D_0 for the melting of a polyphase aggregate in which two or more phases melt simultaneously. If different phases melt in differing proportions then the relationship becomes
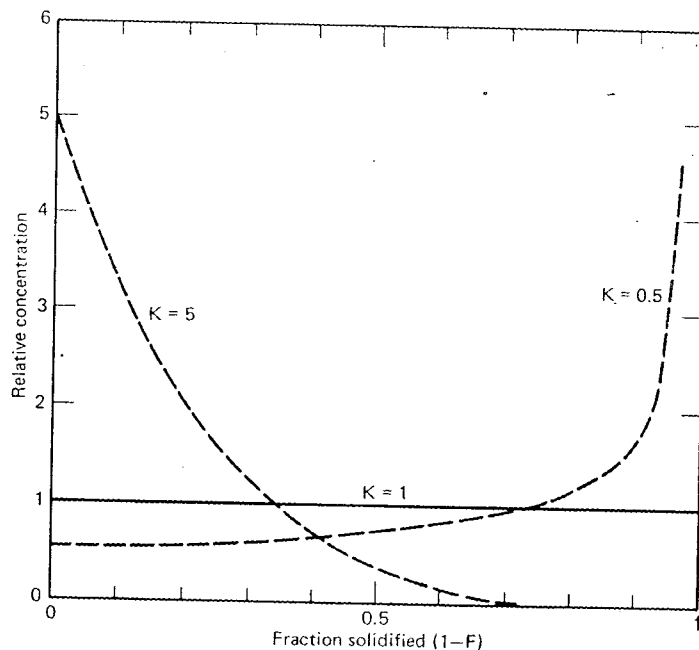
Figure 5.17 Concentration of an element entering the solid phase as a function of fraction solidified

$$C_1 = \frac{1}{D_0} C_0 \left(1 - \frac{PF}{D_0} \right)^{(1/P-1)}$$

where $P = \dfrac{P^\alpha}{K^{1/\alpha}} + \dfrac{P^\beta}{K^{1/\alpha} K^{\alpha/\beta}} + \ldots$

where P^α and P^β are the relative proportions of α and β in the liquid fraction

and $D_0 = \dfrac{X^\alpha}{K^{1/\alpha}} + \dfrac{X^\beta}{K^{1/\alpha} K^{\alpha/\beta}} + \ldots$

where X^α and X^β are the mass fractions of phases α, β. These equations assume that each increment of liquid is removed continuously from preceding liquids. The aggregate liquid composition is given by

$$\overline{C_1} = \frac{C_0}{F} \left[1 - \left(1 - \frac{PF}{D_0} \right)^{1/P} \right]$$

and if the liquid remains at all times in equilibrium the composition C_1 is given by

$$C_1 = \frac{C_0}{D_0 + F(1 - P)}$$

Real problems are complicated by the fact that K and D_0 may not be constant and that incongruent melting may take place. A discussion of such complications is given by Hertogen and Gijbels (1976). Equations of this type have been used to consider

the problem of whether various basalt types can be formed by the partial melting of homogeneous mantle material. It is again important to recall that most melting proceeds at eutectoids and hence simultaneous equations must be used. Even if one solid is not contributing the trace element itself, fusion is providing material to the total liquid. If only a small portion of the source rock is melted to produce magmas then the system is simplified and in some cases the fraction melted may be estimated.

RESIDUAL SOLUTIONS AND PEGMATITES

The data in the preceding sections indicate that the residual melt from fractional crystallization of a magma will in general be a siliceous liquid rich in the alkalies and alumina, containing water and other volatiles, and with a concentration of those minor elements that are not incorporated in the lattices of the common minerals of igneous rocks. Such a residual liquid will probably be unusually fluid for a silicate melt on account of the concentration of volatiles. The pressure of volatiles will also provide the driving force to inject it along surfaces of weakness in the surrounding rocks, which may be the already solid part of the same igneous intrusion or the country rock. In this way pegmatites and some hydrothermal veins may be formed. The literature on pegmatites is very extensive and has been critically summarized by Jahns (1956).

Pegmatites are commonly found in association with granites, as is to be expected if granites are the end product of fractional crystallization of a magma. Granite pegmatites consist essentially of quartz and alkali feldspar, generally with muscovite and biotite also, and are thus similar in composition to granite; the essential difference is in texture, the minerals of pegmatites being of exceedingly variable grain size. Pegmatites are typically coarse-grained, and some of the largest crystals known have been found in them. This is evidently due to low viscosity and concentration of volatiles, which allow the formation of large crystals. Low-viscosity, volatile-rich solutions permit the rapid diffusion of elements and complexes necessary for large crystal growth. Jahns and Burnham (1969) have proposed that pegmatitic textures form when volatile-rich magmas undergo "second boiling" below the critical point for the cooling solution. Two fluid phases form with the more tenuous one providing space for the growth of crystals from the more viscous one. They propose that if the solution does not pass below the critical point, fine-grained aplite dikes rather than pegmatites may form.

Most pegmatites are chemically and mineralogically simple, consisting only of quartz and feldspar, but the complex ones are spectacular in their content of rare elements and unusual minerals; hence they have been intensively studied. Complex pegmatites are economically important and have been worked for lithium, beryllium, scandium, rubidium, yttrium, zirconium, molybdenum, tin, cesium, the rare earths, niobium, tantalum, tungsten, thorium, uranium, and industrial minerals, such as feldspar, muscovite, phlogopite, tourmaline, and quartz. They are often internally zoned, and a specific rare mineral may be confined to a particular zone; zoned pegmatites commonly have a core that consists almost entirely of massive quartz. The geochemistry of pegmatites is a fascinating subject. The major factors influencing the concentration of certain elements in them have already been outlined. Many interesting features, however, such as the localization of one or a few minor elements

in a particular pegmatite or group of related pegmatites, are as yet little understood. Thus the granite pegmatites in southern Norway are noted for their content of rare earths, and tin, lithium, and the rare alkalies are absent. In the pegmatites of the Black Hills (South Dakota) lithium is particularly abundant, tin is common, and the rare earths are absent. Those of central Texas contain rare-earth minerals but carry no other unusual elements. The pegmatites on the island of Honshu in Japan are characterized by the presence of the rare earths and beryllium. These regional differences may, of course, reflect the original composition of the magmas from which they were derived, but the reasons for such compositional differences are unexplained.

Many pegmatites, especially those present in strongly metamorphosed terrains, may have been formed by the differential fusion (anatexis) of preexisting rocks. If pegmatites are the last fraction of a magma to crystallize, then it is reasonable to expect that the first fraction of rock to liquefy (provided suitable volatiles are present) will also have pegmatitic composition. The temperature of formation of pegmatites (probably 500–700°C) has certainly been reached, in many areas, during regional metamorphism. In part the formation of pegmatites may be due to structural control. Rock environments that do not retain residual solutions under pressure may allow such solutions to be disseminated throughout the country rock causing general alteration or greisen–type rocks.

Pegmatites may contain minor amounts of sulfides and arsenides. Gold has been recorded from some. They show many of the characteristics of hydrothermal veins, and examples of pegmatites passing into sulfide-bearing quartz veins have been described. More frequently, however, hydrothermal veins are younger than cogenetic pegmatites. The relationship of hydrothermal solutions to the final stage in the crystallization of a magma is a subject of much difference of opinion. Some geologists believe that hydrothermal solutions are formed as a separate phase later than the pegmatite stage, and laboratory experiments have been cited as indicating that a residual magma of granitic composition would separate on cooling into two immiscible liquids, one of which may be a hydrothermal type solution. On the other hand, it has been found that during crystallization in systems containing an excess of alkali over that required to combine with alumina and silica to give feldspar, the liquid phase varies continuously in composition from a silicate melt containing a small amount of water to a liquid consisting largely of water (i.e., there is a continuous passage from magmatic melt to hydrothermal solution). However, as Verhoogen (1949) stated in a discussion of the thermodynamics of a magmatic gas phase, " ... the number of factors which are relevant in any magmatic process is much larger than is usually recognized, and there are no simple means of deciding in a general way what a magma would do under the most general circumstances. Each case must be examined separately and the idea must constantly be kept in mind that what is true for one magma under certain circumstances is not necessarily true for another magma under the same circumstances, or for the same magma under different circumstances." Hydrothermal activity is most certainly related to magmatic activity but it is possible that the main contribution of the cooling magma may be the heat energy to drive the hydrothermal solutions and the water and dissolved components may be derived in large part from surrounding rocks.

THE VOLATILE COMPONENTS OF A MAGMA

All magmas contain volatile material in an amount and composition not well known, since no means of directly sampling it have yet been devised. From thermodynamic data Krauskopf (1959) has calculated the compositions of possible magmatic gas systems and shows that observed volcanic gas compositions are explicable on the basis of a primary magma containing H_2O, CO_2, H_2S, and minor amounts of HCl, N_2, and HF. His estimate for the average composition of magmatic gas (in atm) is: H_2O, 1000; CO_2, 50; H_2S, 30; HCl, 10; N_2, 10; HF, 0.1. Major variations in the nature of the sulfur species may take place as the oxygen fugacity is changed; H_2S is the predominant sulfur species in reducing magmas, but SO_2 is more important under oxidizing conditions.

The amount of volatile matter in magmas is also subject to considerable differences of opinion. Here we are once more dependent on indirect evidence, such as observations on volcanoes, petrographic examination and chemical analyses of igneous rocks, and experimental determination of the solubility of water and other volatiles in silicate melts. This evidence is in part contradictory. Some volcanologists have stated that in certain eruptions, such as that of Vesuvius in 1906, the mass of gas evolved outweighed the solid matter ejected; in other eruptions the proportion of gas in the material ejected has been estimated as low as 0.3%. Different volcanoes, and the same volcano at different times, evidently present very dissimilar appearances. The data on the solubility of water in artificial silicate melts have indicated that the limiting solubility of water in a granitic melt is about 15% at normal pressures for granite formation, which indicates that magmas probably contain less than this amount. All igneous rocks contain some volatile matter, mostly water, which can be extracted by heating in vacuo, and Clarke's average of the analyses of fresh igneous rocks shows 1.15% H_2O; this figure may perhaps be taken as a lower limit for the average water content of a magma. Current estimates run from about 0.5 to 8% of water in magmas.

As mentioned above, the composition of the volatile components cannot be directly determined. Once again we have to fall back on analyses of the gases evolved on heating igneous rocks, on observations of materials deposited in fumaroles, and on analyses of gases collected around volcanoes. The last procedure, although most direct, is technically difficult and personally uncomfortable or dangerous. Among the most satisfactory collections yet made are those of the gases evolved from the Kilauea volcano (Table 5.6). They are similar in nature to those from several basaltic sources summarized by White and Waring (1963). Averages of such variable figures cannot have more than qualitative significance; they show that the major constituents of the volatile matter at this place are H_2O, CO_2, and sulfur compounds and that H_2O is dominant. Water is usually the major constituent, generally making up over 80% of the whole by volume; CO_2, H_2S, S, SO_2, HCl, and NH_4Cl are often abundant; and HF, N_2, H_2, CH_4, H_3BO_3, and CO have been recorded in lesser amounts. An indication of the change in equilibrium composition of the molecular species in volcanic gas with temperature is shown in Figure 5.18. In addition, gaseous emanations from magmas collect, transport, and deposit many other elements. Such minerals as magnetite, hematite, molybdenite, pyrite, realgar, galena, sphalerite, covel-

Table 5.6   Analytical Data on Kilauea Gases

	Minimum and Maximum Mole Percentages as Collected		Minimum and Maximum Mole Percentages, Recalculated to Exclude O_2, N_2, and Ar	
H_2O	4.7	—99.8		
H_2	0.0	— 0.43	97.2	—99.1
CO_2	0.00023	— 3.8	0.00001	— 0.43
CO	0.0	— 0.086	0.80	— 2.30
CH_4	0.0	— 0.047	0.0	— 0.087
H_2S	0.0	— 0.12	0.0	— 0.023
SO_2	0.0	— 0.16	0.0004	— 0.67
CS_2	0.0	— 0.026	0.001	— 0.43
O_2	0.0	—21	0.0	— 0.026
N_2	0.0	—78		
Ar	0.0	— 0.064		

Heald, Naughton, and Barnes, J. *Geophys. Res.* **68**, 546, 1963.
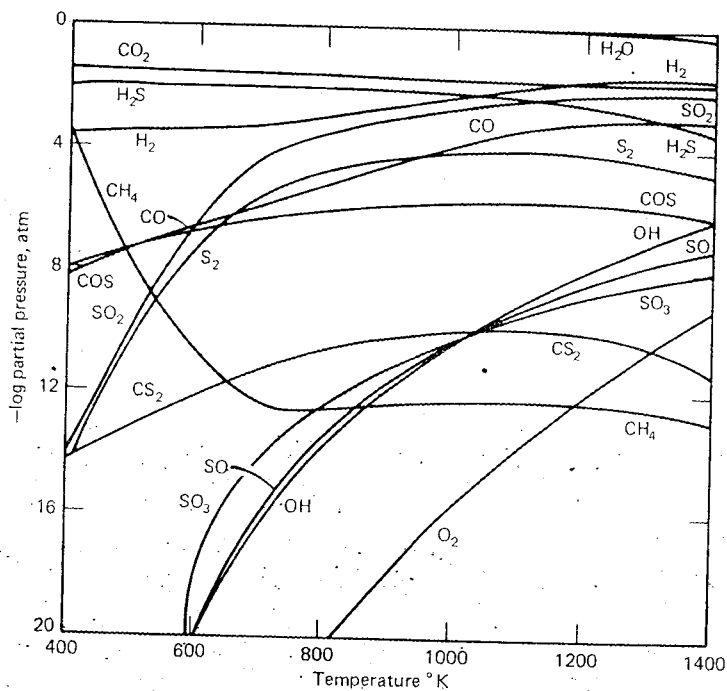
Figure 5.18   Calculated change in equilibrium composition of a volcanic gas phase with change in temperature (1 atm pressure) based on sample taken from Kilauea Iki crater at 1025 K. (From: *Heald, Naughton and Barnes,* see Table 5.6)

lite, sal ammoniac and ferric chloride have been found in the throats of fumaroles where they have been deposited directly from gases. The Valley of Ten Thousand Smokes, Alaska provides us with one of the best-documented examples of the role of volcanic emanations. The work of Allen and Zies showed that the principal gas in all the fumaroles was water-vapor, the percentage by volume ranging from 98.8 to 99.99. Other gases recognized were CO_2, O_2, CO, CH_4, H_2S, H_2, N_2, HCl, and HF. Although these were present in very small percentages, the total amount evolved was enormous. Thus the average concentration of HCl was 0.117% and of HF, 0.032%, but the total quantity of HCl evolved in 1919 was estimated as 1,250,000 tons and of HF, 200,000 tons. One difficulty in discussing the geochemical significance of these figures is that of determining how much of the emanations is primary, that is, was contained in the original magma, and how much secondary, that is, assimilated by the magma from surrounding rocks during its rise to the surface. Allen and Zies believe that much of the water-vapor given off at the Valley of Ten Thousand Smokes was ground water heated by the igneous material. The same difficulty arises in assessing the amount of material in mineral springs that is magmatic in origin. Day and Allen suggest that perhaps 90% of the water in the Yellowstone geysers and hot springs is activated ground water. Precise information on these points would greatly advance our knowledge of the geochemical significance of magmatic emanations, but it is evident that igneous activity during geological time has brought vast amounts of water and other volatile matter to the earth's surface.

MAGMATISM AND ORE DEPOSITION

Much evidence exists for the belief that many ore deposits are genetically related to magmas. Some of this evidence lies in the geological association of ore bodies with igneous rocks, and often with particular types of igneous rocks, for example, tin deposits with granites; sometimes gradations can be traced from pegmatites to ore-bearing veins to barren quartz veins, and examples of direct segregation of ore minerals from magmas are well known. Nevertheless, a significant difference of opinion exists as to the processes whereby ore materials are extracted from their original source, transported, and deposited. Did all or some of the ore-forming material leave a magma as a gas phase or a liquid phase? Was ore the magma's prime contribution or was it heat to mobilize waters from surrounding rocks or to cycle meteoric water? Was the initial ore-forming fluid acid or alkaline? What were the conditions of temperature and pressure? All these questions and many others remain to be answered.

The large majority of mineral deposits appear to have been formed from hydrothermal or hot water solutions, but this does not eliminate the possibility that some deposits, especially those closely connected to magmatic activity, may also have had a gas phase transfer of materials in their genesis. Hydrothermal solutions may be derived in a variety of ways, and it is reasonable to infer that a given solution may have had contributions from magmatic waters, metamorphic waters, connate waters, and ground waters. Oxygen and hydrogen isotope studies shows that ground waters of meteoric origin make a major contribution to most hydrothermal ore deposit systems and that such waters equilibrate rapidly with surrounding rocks. Such water

Table 5.7 Chemical Analysis (in parts per million) of Brine from the Salton Sea Geothermal Field

ph (25°C)	5.2
SiO_2	400
Al	4.2
Fe	2290
Mn	1400
Ca	28,000
Mg	54
Sr	400
Ba	235
Na	50,400
K	17,500
Li	215
Rb	135
Cs	14
NH_4	409
B	390
HCO_3	>150
SO_4	5
Cl	155,000
F	15
Br	120
I	18
NO_3	—
As	12
Cu	8
Zn	540
Pb	102
Ag	1
Total sulfide as H_2S	16

From Muffler and White (1969) *Geol. Soc. Am. Bulletin* 80, 161.

can carry down surface-dissolved material and remove soluble materials from sediments and volcanic rocks they pass through. Heat-mobilized connate waters are related except that they have been in contact with sediments for longer periods of time, may be of marine origin, and have undergone earlier heating cycles. Metamorphic waters may be generated in the conversion of water-rich sedimentary rocks to their metamorphic equivalents. Such waters will dissolve those elements not readily incorporated into the common silicate and oxide rock-forming minerals. Residual water-rich fluids from magmatic differentiation may form ore deposits close to the magma interface or mix with hydrothermal waters from other sources.

Irrespective of their source, hydrothermal solutions appear to be concentrated solutions or brines containing up to 50 wt. % of dissolved solids. Evidence from geothermal brines and fluid inclusions in minerals from ore deposits indicate that the solutions are chlorides of Na, Ca, and K. The composition of the Salton Sea geothermal brine given in Table 5.7 is selected as a typical model.

Other brines and fluid inclusions show local variations in the scarce metals but in general they are found in the range of 1 to 100 ppm. The pH of the solutions may be acid or alkaline but they are generally near to neutral water. From the study of natural solutions and experimental considerations it is estimated that most hydrothermal solutions range from 50 to 550°C, have maximum pressures of 2000 atm, and are high in chloride ion.

High chloride concentration appears to provide the mechanism for the solution and transport of the metals in hydrothermal solutions. Sulfide ion does not appear to be in high enough concentration to complex the metals even though they are most often eventually precipitated as sulfides and hot high-chloride brines will dissolve metallic sulfides. Gold will dissolve by a reaction of the type $Au + 3/2\,O_2 + 6HCl \rightarrow H_3AuCl_6 + 3H_2O$, so that at 500°C and 0.2 kb pressure it has a solubility of over 800 ppm in a $2M$ KCl solution.

The important problem remaining is how the complexed metals are precipitated as sulfides or other compounds. Proposed mechanisms include temperature and pressure changes that could affect the complex equilibrium and reduce the solubility. Chemical reactions of the hydrothermal solutions with wallrocks appear to be a major process in some ore deposits. This is especially evident when wallrocks can provide sulfide ions, or dissolving carbonates produce changes in the hydrothermal solutions. Dilution effects produced by mixing concentrated brines with ground waters provide a probable mechanism for the precipitation of some sulfides.

We are once again faced with the crux of so many geological problems; in general we can observe only the final product and not the steps by which it was formed. No certain criteria have yet been found to provide an unequivocal answer to the questions posed at the beginning of this discussion. The answers undoubtedly vary from one ore body to another and probably from place to place within the same ore body. Seldom can any one of the numerous determining factors be singled out as of unique importance; the question is generally to what degree each has been significant.

SELECTED REFERENCES

Ahrens, L. H. (1964). The significance of the chemical bond for controlling the geochemical distribution of the elements. *Phys. Chem. Earth* **5**, 1–54. An extensive study of the effect of varying degrees of covalency on the distribution of the elements in rocks and minerals.

Barnes, H. L. (ed.) (1979). *Geochemistry of hydrothermal ore deposits (Second Ed.)* 798 pp. John Wiley & Sons, New York. A symposium volume correlating the results of laboratory investigations on the nature of ore-forming fluids with the geochemistry of ore deposits.

Burnham, C. W. (1975) Water and magmas; a mixing model. *Geochim. Cosmochim. Acta* **39**, 1077–1084. Presents a model for the interaction of H_2O with silicate bonds.

Burns, R. G. (1970). *Mineralogical applications of crystal field theory.* 224 pp. Cambridge University Press, Cambridge, England. A summary of crystal field theory and its application to the geochemistry of the transition elements.

Carmichael, I. S. E., F. J. Turner, and J. Verhoogen (1974). *Igneous petrology*. 672 pp. McGraw-Hill Book Co., New York. A comprehensive account of igneous petrology, with careful evaluation of both laboratory and field data.

Deer, W. A., R. A. Howie, and J. Zussman (1962–1963). *Rock-forming minerals*. (5 vols.). 1788 pp. John Wiley & Sons, New York. A comprehensive treatise, with excellent coverage on chemistry, crystal structure, and paragenesis.

Ernst, W. G. (1976). *Petrologic phase equilibria*. 333 pp. W. H. Freeman and Company, San Francisco.

Gast, P. W. (1968). Trace element fractionation and the origin of tholeiitic and alkaline magma types. *Geochim. Cosmochim. Acta* **32**, 1057–1086. A paper on the utilization of trace element data for the elucidation of the genesis of igneous rocks.

Hertogen, J. and R. Gijbels (1976). Calculation of trace element fractionation during partial melting. *Geochim. Cosmochim. Acta*. **40**, 313–322. A review of the complications of trace element fractionation by partial melting. Provides references to earlier important papers.

Hess, H. H. and A. Poldervaart (eds.) (1967). *Basalts*. 862 pp. John Wiley & Sons, New York. A comprehensive account of all aspects of the most abundant group of igneous rocks.

Jahns, R. H. (1956). The study of pegmatites. *Econ. Geol.*, 50th Ann. Vol., 1025–1130. A comprehensive account of the geology and geochemistry of pegmatites; 698 references.

Jahns, R. H. and C. W. Burnham (1969). Experimental studies of pegmatite genesis: A model for the derivation and crystallization of granitic pegmatites. *Econ Geol*. **64**, 843–864.

Krauskopf, K. B. (1959). The use of equilibrium calculations in finding the composition of a magmatic gas phase. In *Researches in geochemistry*, pp. 260–278. John Wiley & Sons, New York. The application of thermodynamic data to the problem of the composition of magmatic gases.

Morey, G. W. (1964). Phase-equilibrium relations of the common rock-forming oxides except water. *U. S. Geol. Surv. Prof. Paper 440-L*, 159 pp. (Data of Geochemistry, sixth ed.; chapter L). A comprehensive account of the phase relations of the common rock-forming oxides and the binary, ternary, quaternary, and quinary systems formed by them.

Nockolds, S. R. and R. Allen (1953–1956). The geochemistry of some igneous rock series. *Geochim. Cosmochim. Acta* **4**, 105–142, **5**, 245–285; **9**, 34–77. A series of papers describing the distribution of minor trace elements in a considerable number of igneous rock series.

Osborn, E. F. (1962). Reaction series for subalkaline igneous rocks based on different oxygen pressure conditions. *Amer. Mineral.*, **47**, 211–226. This paper established the important effect of oxygen pressure in modifying the classic reaction series for the evolution of igneous rocks.

Ringwood, A. E. (1955). The principles governing trace element distribution during magmatic crystallization. *Geochim. Cosmochim. Acta*, **7**, 189–202, 242–254. A discussion of trace element distribution in igneous rocks, emphasizing the importance both of ionic size and of electronegativity.

Shaw, D. M. (1977). Trace element behavior during anatexis. In *Magma genesis.* H. J. B. Dick (ed.), Bulletin 96, Oregon Dept. of Geology and Mineral Industries, 189–213. A discussion of trace element distributions, contains many useful references.

Skinner, B. J. and P. B. Barton (1973). Genesis of mineral deposits. *Ann. Rev. Earth Planet. Sci.* **1**, 183–211. A comprehensive review article on recent advances in the field of ore genesis; extensive bibliography.

Taylor, S. R. (1965). The application of trace element data to problems in petrology. *Phys. Chem. Earth* **6**, 133–213. An extensive review article on the interpretation of trace element abundances in rocks and minerals; excellent bibliography.

Verhoogen, J., (1949). Thermodynamics of a magmatic gas phase. *Univ. Calif. Pub. Bull. Dept. Geol. Sci* **28**, 91–136. A precise discussion of the factors governing equilibrium in magmatic processes.

Wager, L. R. and G. M. Brown (1967). *Layered igneous rocks.* 588 pp. W. H. Freeman and Co., San Francisco. A comprehensive account of the Skaergaard intrusion and other igneous bodies illustrating magmatic crystallization.

Wager, L. R. and R. L. Mitchell (1951). The distribution of trace elements during strong fractionation of basic magma—a further study of the Skaergaard intrusion, East Greenland. *Geochim. Cosmochim. Acta* **1**, 129–208. An extensive study of the content of minor and trace elements in a strongly differentiated intrusive body.

Yoder, H. S. (1976). *Generation of basaltic magma.* 265 pp. National Academy of Sciences, Washington. An updated review of the many aspects of basalts and their formation.

6 SEDIMENTATION AND SEDIMENTARY ROCKS

SEDIMENTATION AS A GEOCHEMICAL PROCESS

Sedimentation is in effect the interaction of the atmosphere and hydrosphere on the crust of the earth. Various aspects of sedimentation are described in terms of weathering, erosion, deposition, and diagenesis, but no one of these processes works in isolation. The original constituents of the crust, the minerals of igneous rocks, are to a large extent unstable with respect to the atmosphere and hydrosphere. They have been formed at high temperatures, and sometimes at high pressures also, and cannot be expected to remain stable under the very different conditions at the earth's surface. Of the common minerals of igneous rocks, only quartz is highly resistant to weathering processes. All the other minerals tend to alter; by the action of oxygen, carbonic acid, and water, they are more or less attacked, and new minerals are formed which are more stable under the new conditions. The altered rock rapidly crumbles under the mechanical effects of erosion, and its constituents are transported by wind, water, or ice and redeposited as sediments or remain in solution.

The key reactions in the geochemistry of sedimentation are the chemical breakdown of some minerals and the formation of others. Of these, silicates are the most important, since they constitute more than 90% of the earth's crust (including quartz as a silicate). The processes by which silicate minerals are broken down chemically during weathering have long been a subject of speculation. It was early found that the univalent and bivalent cations readily go into solution, but the fate of aluminum and silicon has been less well understood. Some hypotheses assumed the hydrolysis of aluminosilicates with the formation of colloidal silicic acid and aluminum hydroxide, which later reacted to give clay minerals. Other investigations have shown, however, that during the initial attack some silicate minerals go into ionic solution, and

even the silica and the alumina are at least for a short time in true solution. At the surface of a crystal unsatisfied valencies exist that are the loci of reaction with water molecules. Hydration and hydrolysis follow, whereby strong bases, such as potassium, calcium, and magnesium, are removed and oxygen anions in the crystal lattice may be partly replaced by hydroxyl ions. Aluminum and silicon attract OH ions strongly; aluminum probably groups six OH around it to assume its preferred sixfold coordination, whereas silicon remains four-coordinated. When first set free these elements are in ionic solution, but the ions tend to aggregate and to form clusters of collodial size. When first formed these colloidal aggregates are probably amorphous, but, on ageing, orientation into definite crystal lattices, such as those of the clay minerals, takes place. Some silicate minerals may not undergo complete breakdown of the lattice during weathering; for example, biotite and muscovite may perhaps pass directly into clay minerals by ionic substitution, whereby fragments of the sheet structures may be directly incorporated in the new minerals. Likewise, the persistence of aluminum in tetrahedral coordination, together with the retention of mobile cations in some clay minerals, is often interpreted as evidence that part of parent feldspar structure is sometimes retained by weathering products. The ultimate fate of different elements thus depends largely on the relative stability of their ions in water. The most stable are the alkali metal ions, followed by those of the alkaline earths, and they are for the most part carried away in solution. Silicon, aluminum, and iron, on the other hand, are generally soon redeposited as insoluble compounds; new minerals are formed from them at an early stage of weathering. The specific weathering reactions may thus be characterized as either congruent, in which the weathered material is removed, or incongruent, in which it is in part converted to other relatively immobile mineral species. Included in the congruent type reactions are simple solution by water, water acidified by carbon dioxide, or by chelation by organic molecules in solution. Water in contact with the atmosphere contains $10^{-5}M$ H_2CO_3, but carbonic acid is often greatly concentrated in soil waters by the decay and oxidation of organic material. Incongruent reactions include hydration and carbonation to form insoluble hydroxides and carbonates, oxidation, most notably of

Table 6.1   Chemical Data for the Weathered Sequence Developed on Basalt at Bathurst, New South Wales[a]

No.	Depth	Analysis (%)									
		SiO_2	Al_2O_3	Fe_2O_3	CaO	MgO	Na_2O	K_2O	TiO_2	H_2O	Total
B1	15 cm	39.3	20.3	16.4	1.7	0.5	0.3	0.6	2.4	18.6	100.1
B2	46 cm	36.1	20.6	21.8	0.5	0.3	0.2	0.2	4.1	16.3	100.1
B3	76 cm	27.8	18.5	33.2	0.5	0.2	0.2	0.1	5.2	14.3	100.0
B4	107 cm	28.9	19.4	30.5	0.5	0.2	0.2	0.1	4.8	15.8	100.4
B5	137 cm	33.4	21.4	24.3	0.5	0.3	0.2	0.1	4.1	15.8	100.1
B6	168 cm	37.1	23.2	18.1	0.6	0.3	0.2	0.1	3.8	16.8	100.2
B7	198 cm	42.2	15.2	14.0	9.0	8.7	1.1	0.9	2.0	7.2	100.3
B8	213 cm	45.0	14.6	13.3	9.8	10.1	2.4	1.5	1.9	1.6	100.2

[a]After Loughnan (1969).

ferrous iron compounds to form insoluble ferric iron oxides or hydroxides, and the weathering of silicates discussed above.

Irrespective of the specific weathering mechanisms, the chemical changes due to weathering are directly evident if chemical analyses are made through a weathered sequence of rock material. Such analyses have been made on different parent rock types in different climates. Typical data are shown in Table 6.1. The rapid loss of soluble cations from the weathered basalt and the increase in insoluble oxides may readily be seen. Figure 6.1 illustrates the change in mineralogy accompanying the chemical weathering. The sequence illustrated is typical of temperate climates. Weathering profiles in tropical regions approach alumina and at times titania-rich residues due to the leaching of SiO_2 aided by the high concentration of organic acids in soil waters. Some residues show increases in particular elements due to accumulation of particularly resistant minerals such as ilmenite, rutile, zircon, quartz, or even gold and platinum.

SOIL GEOCHEMISTRY

Weathering and sedimentation are the geochemical processes of greatest importance to humans, since they provide us with our basic economic resource, the soil. Human culture and civilization can be closely correlated with the pattern of soil fertility. This can be directly traced back to the geochemical processes that have been responsible for the formation of the soil from the parent rock materials.

The unique characteristic of soil is the organization of its constituents and properties into layers that are related to the present-day surface and change vertically with depth. The individual layers are referred to as soil horizons and may range in thickness from a few inches to several feet; collectively, they are known as the soil profile. Most soil profiles comprise three principal horizons, identified from the surface downward as A, B, and C. The A horizon develops primarily as a result of partial loss of original material by leaching and mechanical removal (eluviation), resulting from the downward percolation of rain water. This eluviated material accumulates in the B horizon, which is thus a zone of accumulation. The B horizon is characteristically enriched in clay and often has a red-brown or yellow-brown color from an accumulation of iron oxides. Minor and trace elements are frequently enriched in the B horizon. The C horizon is the parent material for the overlying A and B horizons; it may be rock in situ, transported alluvial or glacial material, or even the soil of a preceding cycle.

Because soil results from the weathering of rocks, its composition must depend on the composition of the rock from which it was formed. This statement is obvious, but it may be misleading. Although soils do differ in composition, they are on the whole rather uniform in a given area, and the differences are primarily the result of environmental factors. The same parent rock may give rise to very different soils under different conditions. The environmental factors include climate, biological activity, topography, and time, and the most important of these is climate. This can be seen by comparing the productivity of soils formed from the same rock type in the normal temperate zone and in the humid tropics. In general, the soil in the humid tropics will be much less fertile, as a result of the intense leaching brought about by high
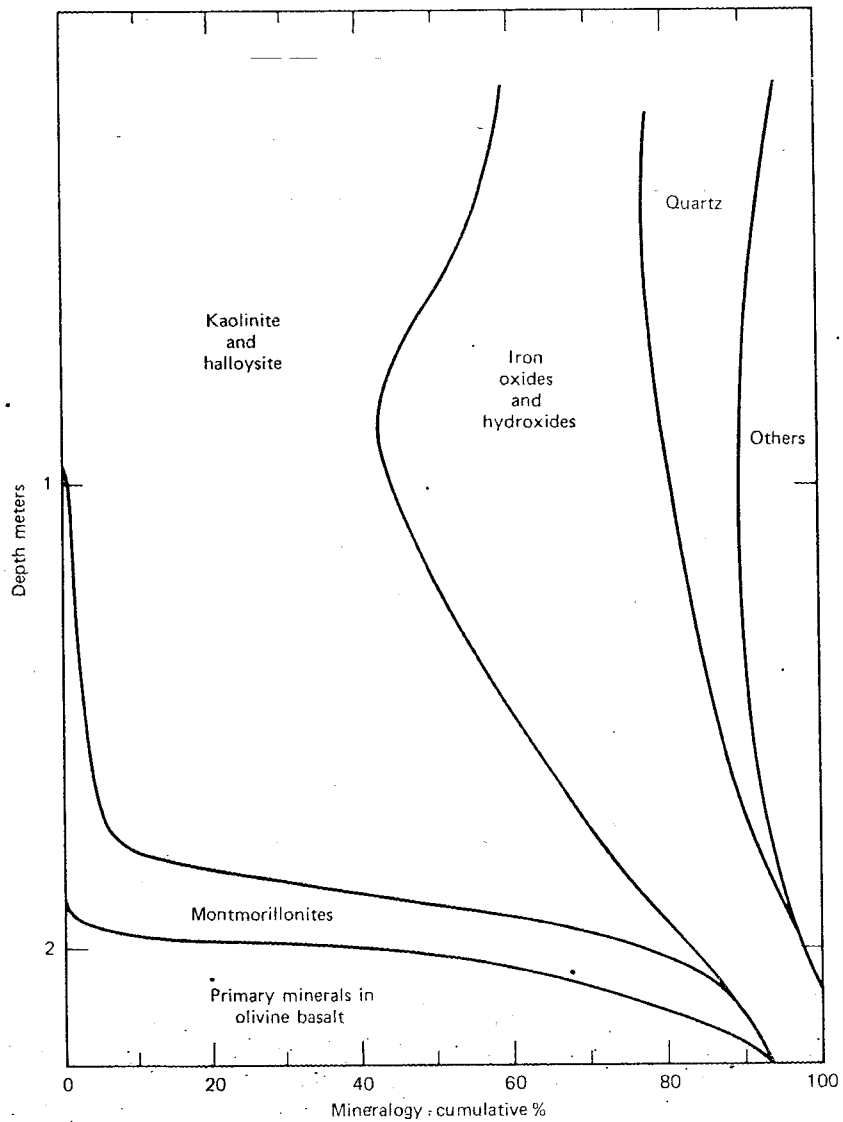
rainfall, high temperature, and the almost complete removal of organic matter by microorganisms.

The components of a soil include not only the minerals present, but also organic matter, water, and air. These components are generally not in equilibrium with the environmental conditions, and are hence in a continual state of change. The transformation from parent rock to soil is generally accompanied by a marked decrease in Ca, Mg, Na, and K, relatively smaller losses of Al and Fe, and increase in Si. The most active part of the soil is the colloidal fraction, which consists mostly of clay minerals, together with organic matter. Many, although not all, of the important agronomic characteristics of soils depend on the quantity and kind of colloids present and their state of aggregation. True soil cannot be formed without the presence of some organic matter. Mere chemical and physical weathering of rocks does not necessarily result in soil formation, as is shown by the absence of soils in true deserts, either tropical or arctic. Most soil processes are directly or indirectly biological in nature. Organisms are effective agents for extracting and dissolving many elements. Because of the enormous multiplication rate of microorganisms, their total effect can be considerable and is probably significant in the migration of minor and trace elements in soils.

The data on major and minor elements in the parent rock can be used as an indication of the background composition of the soil. However, background composition in soils is also subject to considerable variation according to soil type and soil horizon, particularly in well-differentiated profiles characterized by marked enrichment in iron oxides or organic matter. The range of values for some minor and trace elements in normal soils is given in Figure 6.2.
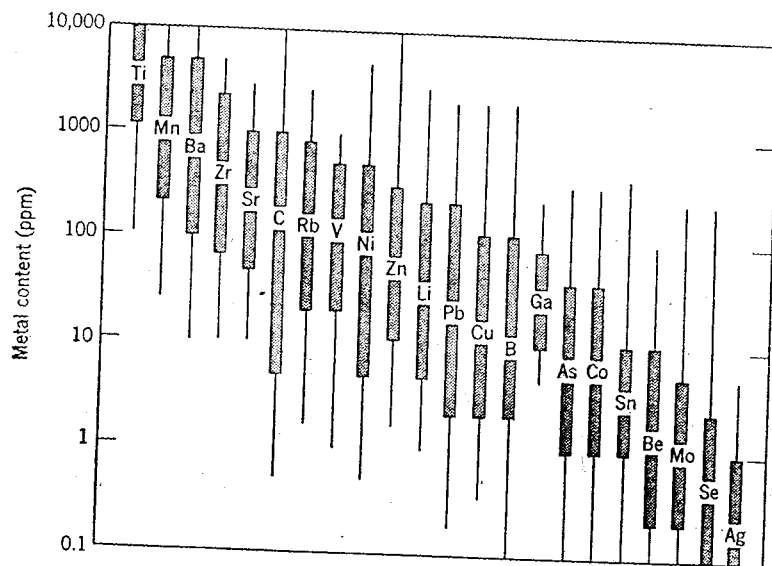
Figure 6.2 Range of the content of some minor elements in normal soils. Thin lines indicate more unusual values. (From Mitchell, in Bear, *Chemistry of the soil.* Courtesy of Reinhold Publishing Corp.)

The importance of trace elements in the nutrition of plants, and through plants of animals, has been increasingly realized in recent years. Good crops and healthy humans and domestic animals are closely connected with the presence in the soil of a number of minor and trace elements. The productivity of many regions has been greatly enhanced by the application of certain elements in minute amounts. This subject is essentially part of soil geochemistry, because plants derive these trace elements from the soil, not directly from the rocks. It is interesting to note that some trace elements essential for animal health seem not to be required by plants. Grass grows on soil deficient in cobalt as well as on soil with a normal cobalt content, but animals fed on such grass develop a deficiency disease that can be corrected by the feeding of trace amounts of cobalt, either directly or through the addition of cobalt-containing fertilizers to the soil. The possible correlation of the geographical distribution of disease and soil geochemistry is a field of extreme significance that as yet has been inadequately studied. In part this is because of the extreme complexity of the problem and the difficulty of isolating the numerous factors involved.

THE CHEMICAL COMPOSITION OF SEDIMENTARY ROCKS

The chemical composition of sedimentary rocks is exceedingly variable, more so than that of igneous rocks, since sedimentation generally leads to a further diversification. Considering compositions in terms of oxides, we find that SiO_2 may exceed 99% in some sandstones; Al_2O_3 may reach nearly 70% in bauxite; Fe_2O_3 up to 75% in limonite; FeO as high as 60% in siderite; MgO to 20% in dolomite; and CaO to 56% in pure limestones. In view of such variations, the determination of the average chemical composition of sedimentary rocks is not simple. The method used by Clarke and Washington for igneous rocks is hardly applicable because of a lack of analyses of sedimentary rocks and of inadequate sampling. There is little urge to analyze these rocks unless they have economic significance, and then they are generally of unusual composition.

Clarke estimated the average composition of the common sedimentary rocks— shale, sandstone, and limestone—by analyzing mixtures of many individual samples (Table 6.2). Then, by using an estimate for the relative amounts of these common sedimentary rocks, an average for sedimentary rocks as a whole can be calculated. Such an average is given in Table 6.2 using the figures shale 80%, sandstone 15%, and limestone 5%, given by Clarke. Garrels and Mackenzie (1971) have made an extensive analysis of this problem, using more recent data and basing their calculation on the geochemical balance involved in the conversion of the average igneous rock into the average sedimentary rock. They calculate the proportion of shale : sandstone : limestone to be 81 : 11 : 8, and they arrive at an average composition for sedimentary rocks not significantly different from that obtained by Clarke. However, there is some doubt as to the acceptability of these figures, since they seem to understate the amount of limestone. Ronov and Yaroshevsky (1969) calculated an average composition of all sediments, based on what is known of their distribution, mass, and individual compositions. When their figures are compared with the preceding estimates (Table 6.2) the differences are clearly due largely to the greater amount of calcium carbonate allowed for by Ronov and Yaroshevsky.

Table 6.2 Chemical Composition of Sedimentary Rocks

	Average Igneous Rock	Average Shale	Average Sandstone	Average Limestone	Average Sediment*	Average Sediment†	Average Sediment‡
SiO_2	59.14	58.10	78.33	5.19	58.49	59.7	46.20
TiO_2	1.05	0.65	0.25	0.06	0.56	—	0.58
Al_2O_3	15.34	15.40	4.77	0.81	13.08	14.6	10.50
Fe_2O_3	3.08	4.02	1.07	0.54	3.41	3.5	3.32
FeO	3.80	2.45	0.30	—	2.01	2.6	1.95
MgO	3.49	2.44	1.16	7.89	2.51	2.6	2.87
CaO	5.08	3.11	5.50	42.57	5.45	4.8	14.00
Na_2O	3.84	1.30	0.45	0.05	1.11	0.9	1.17
K_2O	3.13	3.24	1.31	0.33	2.81	3.2	2.07
H_2O	1.15	5.00	1.63	0.77	4.28	3.4	3.85
P_2O_5	0.30	0.17	0.08	0.04	0.15	—	0.13
CO_2	0.10	2.63	5.03	41.54	4.93	4.7	12.10
SO_3	—	0.64	0.07	0.05	0.52	—	0.50
BaO	0.06	0.05	0.05	—	0.05	—	—
C	—	0.80	—	—	0.64	—	0.49
	99.56	100.00	100.00	99.84	100.00	100.0	100.13

*Shale 80, sandstone 15, limestone 5; after Clarke.
†Garrels and Mackenzie, 1971.
‡Ronov and Yaroshevsky, 1969; includes MnO 0.16, Cl 0.24.

It has been considered that the average composition of sedimentary rocks should correspond fairly closely to that of igneous rocks, since all sedimentary rocks have ultimately been derived from igneous rocks by processes of weathering. The only permanent change should be the loss of those elements, principally sodium, that tend to accumulate in solution in the ocean, and the addition of components from the atmosphere and hydrosphere, such as oxygen, carbon dioxide, and water. However, if the figures of Ronov and Yaroshevsky are reliable, the average sediment does not correspond to the average igneous rock. This apparent discrepancy in the geochemical balance sheet, especially for calcium, is a significant problem. It may be related, in part at least, to a marked increase of pelagic foraminifers in Tertiary and Recent times, leading to a greater precipitation of calcium carbonate in young sediments and sedimentary rocks.

Significant features of the chemical composition of sedimentary rocks are the dominance of potassium over sodium, alumina in excess of the 1:1 mole ratio to alkalies and calcium, high silica in sandy and cherty rocks, high lime and magnesia in the carbonates, and the presence of iron mainly in the ferric state. These are generalizations, and many exceptions may be found. The relationship between chemical composition and rock type in the sediments can be expressed in a gross fashion by a triangular composition diagram in which the apices are SiO_2, $(Al, Fe)_2O_3 xH_2O$, and $(Ca, Mg)CO_3$ (Figure 6.3). Such a diagram neglects the alkalies, but they are generally low except in some argillaceous rocks. It is difficult, if not impossible, to define sedimentary rock types from chemical analyses alone by any such system as the norm for igneous rocks. The main groups overlap, as Figure 6.3 shows. Some tentative
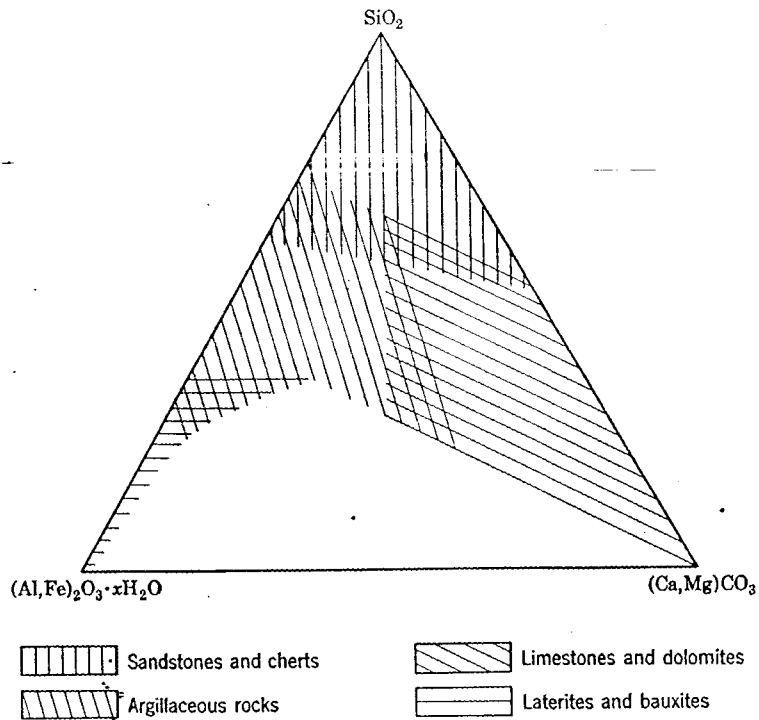
Figure 6.3 The chemical composition of common sediments. Sediments with compositions falling in the blank area are rare or nonexistent.
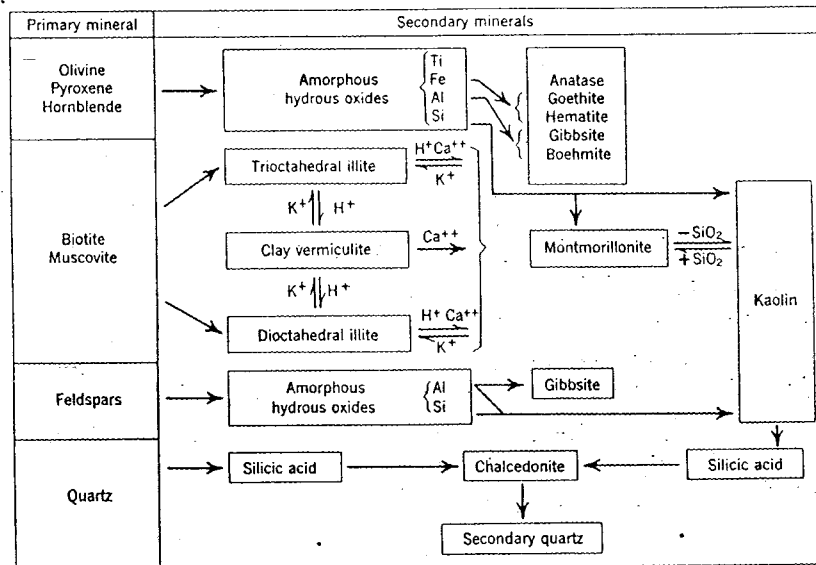
Figure 6.4 Weathering of primary rock-forming minerals. (After Fieldes and Swindale, *New Zealand J. Sci. Tech.* **36B**, 140, 1954)

limits can be stated: argillaceous rocks with more than 50% SiO_2 generally contain free silica, whereas those with more than 40% Al_2O_3 contain free alumina. To call a rock a limestone or dolomite should at least imply that carbonate is the dominant component, and similarly in sandstone or chert free silica should exceed any other component. Imbrie and Poldervaart (1959) have developed a procedure for the calculation of mineralogical compositions from chemical analyses of sedimentary rocks.

THE MINERALOGICAL COMPOSITION OF SEDIMENTARY ROCKS

The mineralogy of sedimentary rocks is characterized by two distinct types of material: resistant minerals from the mechanical breakdown of the parent rocks, and minerals newly formed from the products of chemical decomposition. The latter minerals are generally hydrated compounds, as is to be expected in substances formed in a water-rich environment. Goldich (1938) pointed out that the order of stability of minerals of igneous rocks toward weathering is the reverse of their order in the reaction series of Bowen; thus

Stability decreases

$\Biggl\downarrow$

Quartz	
Muscovite	
Potash feldspar	
Biotite	
Alkalic plagioclase	
Alkali-calcic plagioclase	Hornblende
Calcic-alkalic plagioclase	
	Augite
Calcic plagioclase	
	Olivine
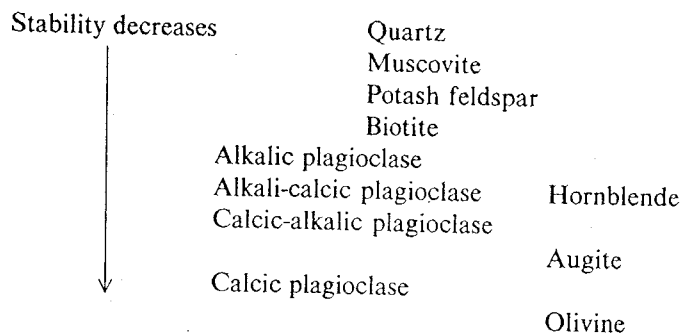
The arrangement does not of course imply a reverse reaction series; the minerals do not invert one into the other on weathering but are decomposed into their components. This identity of arrangement between Bowen's reaction series and Goldich's stability series indicates that the last-formed minerals of igneous rocks are more stable at ordinary temperatures than are the minerals formed at an early stage of crystallization. In other words, the differential between conditions at the time of formation and those existing at the surface reflects the order of stability of the common silicates of igneous rocks. Figure 6.4 presents a scheme that shows how primary minerals change to successive secondary minerals through weathering processes. The primary minerals are listed in order of increasing resistance to weathering, and the secondary minerals are arranged to indicate their relation to the minerals from which they originate.

The total number of minerals recorded from sedimentary rocks is very large, because almost any mineral of igneous or metamorphic origin may have at least a transitory existence in a sediment. Nevertheless, the common and abundant minerals of sedimentary rocks are few: quartz, feldspar, calcite, dolomite, and clay minerals. Some other minerals, such as glauconite, limonite (goethite and hematite), bauxite

(gibbsite and boehmite), and collophane (sedimentary apatite), are abundant in sediments of restricted extent, and amorphous colloidal material, such as allophane, may be present.

Quartz and the feldspars are the abundant detrital minerals, that is, those set free by the mechanical breakdown of preexisting rocks. Quartz is very resistant to chemical attack under surface conditions. The feldspars are less resistant, for, although they may persist indefinitely in sedimentary rocks, they are chemically decomposed by prolonged weathering. On the other hand, the alkali feldspars may undoubtedly be formed in sediments under surface or near-surface conditions, since authigenic albite, orthoclase, and microcline have been recorded. This variability in behavior presumably reflects the environment; solution probably takes place in acid environments, whereas in alkaline environments authigenic feldspar can form. Anorthite is readily decomposed by weathering, but it is interesting to note that authigenic calcium zeolites (heulandite, chabazite, laumontite, and others) have been found in some sediments and sedimentary rocks.

Calcite is precipitated from solution either by physicochemical changes or by the vital processes of organisms. Aragonite is sometimes the form in which calcium carbonate is deposited, but in general it does not persist, for it inverts more or less readily into calcite, the more stable form. The origin of dolomite has been a subject of much discussion, and general agreement has not been reached. However, geological evidence indicates that many dolomites have been formed from limestones by the metasomatic action of magnesium-bearing waters. In many instances sea water acting on calcium carbonate during diagenesis has evidently been responsible. For the reaction

$$2CaCO_3 + Mg^{2+} = CaMg(CO_3)_2 + Ca^{2+}$$
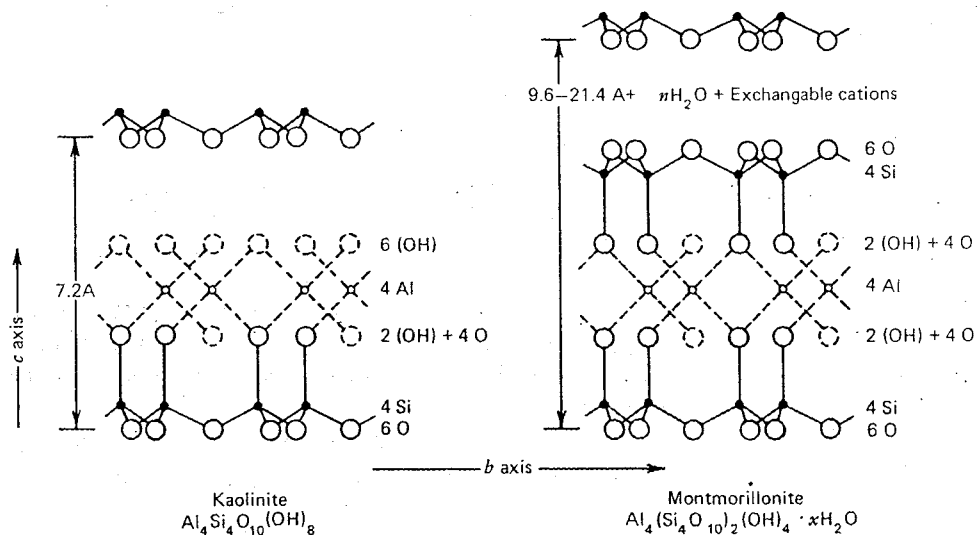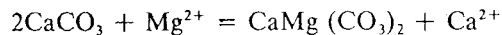
Figure 6.5 Schematic presentation of the structures of the principal clay minerals. (After Grim, *J. Geol.* **50**, 225–275, 1942)

the law of mass action predicts that equilibrium will be determined almost entirely by the relative concentration of calcium and magnesium ions in solution. Studies on the thermodynamics of the reaction have shown that with the conditions of temperature and concentration prevailing in the sea the free-energy change of the above reaction is negative, that is, dolomitization will proceed spontaneously. Experimental studies of dolomite genesis are very difficult, due to the very slow rate of cation ordering in the dolomite lattice. In natural systems sufficient time for the slow ordering reactions is available.

Kaolinite, montmorillonite, illite, and chlorite, together with a number of less common species, make up the clay minerals of sediments and sedimentary rocks. Clay mineralogy has been the subject of intensive research in recent years, and a detailed account has been given by Grim (1968). The clay minerals are stable secondary products formed by the decomposition of other aluminosilicates. It is significant that all these new-formed minerals have layer-lattice structures, which seem to have greater stability than other types under surface conditions. Besides their characteristic crystal structure, the clay minerals have other features in common. They are all hydrous aluminosilicates. In most sediments their average grain size is very small, generally less than 0.002 mm in diameter and ranging down to colloidal dimensions; as grain size diminishes so does apparent perfection of crystallinity, and probably no sharp break exists between imperfectly crystallized clay minerals and amorphous material, which may perhaps be regarded as alumino-silicate gel. Very often more than one clay mineral is present in a particular sediment. Not only mechanical mixtures, but also "mixed-layer" crystals, in which molecular layers of more than one clay-mineral species are interleaved in a single crystal, have been recognized. Thus the complications are many, and positive identification of the phases in the clay frac-

Muscovite
$K_2Al_4(Si_6Al_2)O_{20}(OH)_4$

Chlorite
$Mg_{10}Al_2(Si_6Al_2)O_{20}(OH)_{16}$
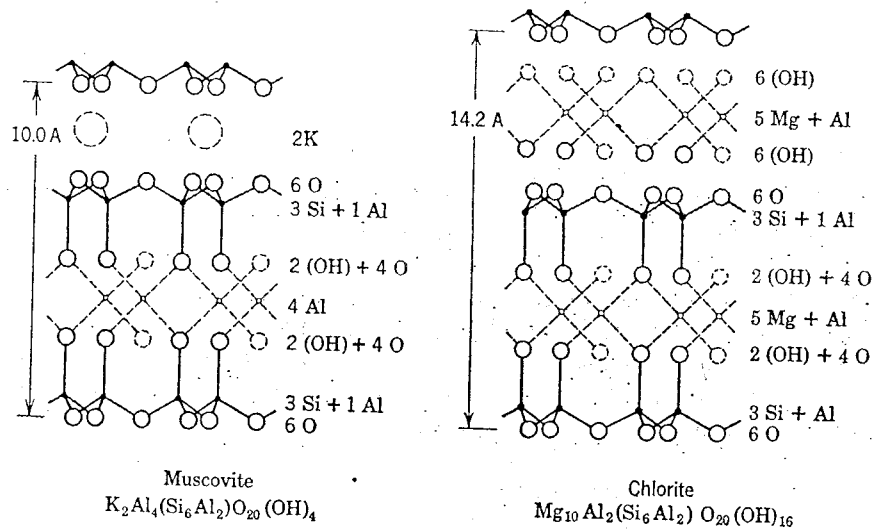
Figure 6.5 (continued)

tion of sediments may be considered the most exacting problem with which a mineralogist is faced.

The similarities and differences between the clay minerals can best be understood by consideration of their structures (Figure 6.5). They are all sheet silicates, and their structures can be considered as produced by the stacking of two different units in the direction of the c axis. These units are (a) tetrahedral sheets of linked (Si_4O_{10}) sheets; and (b) octahedral sheets of aluminum-hydroxyl units, consisting of aluminum ions between two sheets of close-packed hydroxyls or oxygens; each aluminum is surrounded by six oxygens or hydroxyls, that is, is in sixfold coordination. These two types of units are linked in the clay mineral by oxygens common to both. The individual clay minerals differ in the relative number of the two types of units in their structures and in the possibility of replacement of silicon or aluminum by other elements.

The structure of kaolinite consists of one tetrahedral sheet linked with one octahedral sheet, that is, a two-layer structure. Replacement of silicon and aluminum by other elements does not occur, and so analyses of pure kaolinite always correspond closely with the ideal formula. Kaolinite is one of four polymorphs, the others being dickite, nacrite, and halloysite. Dickite and nacrite are generally of hydrothermal origin and rarely occur in sediments. Halloysite occurs in hydrothermal deposits, but it is also found occasionally in sedimentary rocks, where it has generally been precipitated from acid ground water carrying alumina and silica in solution.

The montmorillonite structure has layers consisting of one aluminum-hydroxyl unit sandwiched between two (Si_4O_{10}) sheets; these layers are stacked one above the other in the c direction, with water molecules between them. A characteristic feature of montmorillonite is the variable water content, which is reflected in the c repeat, varying from 9.6 Å in dehydrated material to 21.4 Å when the mineral is water-saturated. Montmorillonite is therefore said to have an expanding lattice; the characteristic swelling properties of bentonites in water are due to their montmorillonite content.

Considerable atomic substitution is possible in the montmorillonite structure; the aluminum can be partly or wholly replaced by ferric iron (nontronite), by magnesium (saponite), by zinc (sauconite), and by smaller amounts of lithium, trivalent chromium, manganese, and nickel; the silicon can be partly replaced by aluminum, giving the variety beidellite. The chemical composition of montmorillonite (in the group sense) is therefore exceedingly variable.

Montmorillonite always differs from the theoretical formula in that the three-layer unit has a net negative charge resulting from the substitutions noted above. Many analyses have shown this to be about two-thirds of a unit per unit cell. This charge deficiency is balanced by exchangeable cations adsorbed between the unit layers and around their edges; these adsorbed cations are often calcium and sodium, elements that are therefore found in analyses of montmorillonite, although the ideal structure has no lattice positions for them.

A common and important constituent in many clays and shales is the material known as illite or hydromica. The occurrence of a micaceous clay material with less potassium and more water than the theoretical composition of muscovite has been

recognized for many years. The name illite is useful as a general term, not as a specific clay-mineral name. Some illite is essentially clay-size muscovite; some is a mixed-layer muscovite-montmorillonite; and some is a mechanical admixture of muscovite and montmorillonite, often with some clay-size quartz.

Chlorite may be an important constituent of argillaceous material. Small amounts of chlorite admixed with other clay minerals are particularly difficult to detect. Structurally, chlorite may be derived from montmorillonite by the insertion of a (Mg,Al)(OH) layer between each montmorillonite layer, just as muscovite is derived by the insertion of a layer of potassium ions. This structural relation may also be a paragenetic relation, since there is evidence for the formation of chlorite from montmorillonite in sea water, which is rich in magnesium ions to provide the (Mg,Al)(OH) layer.

Much remains to be learned of the mutual relations of the clay minerals and of the conditions favoring the formation of one in preference to another. The primary factors determining the nature of a clay are, first, the chemical character of the parent material and, second, the physicochemical environment in which alteration of this material takes place. The structure of kaolinite does not accommodate cations other than silicon and aluminum, and the formation of kaolinite is evidently favored by an acid environment, in which all bases tend to be removed in solution. Kaolinite also has the highest Al:Si ratio of the clay minerals, and its formation is promoted if the weathering processes tend to remove silica in solution, thereby enriching the residue in alumina. This is reflected in the geological evidence for the common formation of kaolinite by the alteration of alkali feldspar in an acid environment, the Na^+ and K^+ ions released tending to stabilize the silica in solution. Montmorillonite, on the other hand, has an Si:Al ratio about 2:1, and other cations, such as magnesium and iron, are probably essential to its formation; it seems to form most readily in a neutral or slightly alkaline environment from ferromagnesian minerals, calcic feldspars, and especially volcanic ash. Montmorillonite has a close structural relationship to illite (muscovite) and chlorite and evidently changes readily to these minerals, especially in the marine environment. The comparatively high concentration of potassium and magnesium in sea water promotes this change. Illite is the commonest clay material in marine sediments and sedimentary rocks. The fixation of potassium in illite and magnesium in chlorite is probably an important mechanism in regulating sea-water composition. In general, diagenesis promotes the formation of illite and chlorite and the disappearance of kaolinite and montmorillonite, so that shales and argillites consist largely of the former minerals.

PHYSICOCHEMICAL FACTORS IN SEDIMENTATION

The geochemistry of sedimentary processes is essentially the geochemistry of reactions taking place in the presence of water. Water is by no means a typical liquid, and in this connection its unique properties deserve to be emphasized. As a solvent water is unequalled; no other liquid can compare with it in the number of substances it can dissolve nor in the amounts it can hold in solution. The structure of the water molecule is the key to these remarkable properties. The hydrogens are small in rela-

tion to the oxygen atom so that the molecule is approximately spherical. Its radius is only slightly greater than that of the oxygen ion. The angle between the bonds joining the hydrogen nuclei to the center of the oxygen atom is 105°. This structure results in a very uneven distribution of charge in the molecule. An excess of positive charge appears at or between the protons, and the opposite side of the molecule is negatively charged. Thus the water molecule is a dipole. The mutual attraction of these dipoles makes the cohesive forces between water molecules very much larger than for normal liquids, which owe their cohesion to weak van der Waals' forces. Hence water, for a liquid of low molecular weight, has extremely high melting and boiling points. The liquid range of water, 100° at 1 atm, is unexpectedly long because of the persistence of bonding between the molecules. The energy necessary to break the remaining bonds at the boiling point is reflected in the abnormally high heat of vaporization, 9720 cal/mole. The cohesive forces between water molecules are also reflected in the surface tension, which is 72.75 dynes/cm at 20°, compared with 26.77 for CCl_4 and 28.88 for C_6H_6. Another important effect of the dipole nature of the molecules is to give water its abnormally high dielectric constant, namely, 80. The high dielectric constant is responsible for its activity as a solvent for ionic compounds, because the force of attraction between ions varies inversely as the dielectric constant of the medium, and solution of ionic compounds is essentially a dispersion of the ions by the molecules of the solvent.

IONIC POTENTIAL

Ions in solution attract water molecules to them; cations attract the negative ends of nearby water dipoles, anions the positive ends. The number of water molecules that can thereby be packed around a given ion depends on the size of the ion; the bigger it is, the greater the number of water molecules that can cluster around it. However, the degree of hydration depends not only on the size of the ion, but also on the intensity of the charge on its surface. For example, the lithium ion, with a radius of 0.74 Å, exerts a much stronger attraction for water dipoles than the cesium ion, with a radius of 1.70 Å, although the charge on each is the same. As a result, the lithium ion is hydrated, despite its small radius, whereas the cesium ion is not. The hydration of an ion is thus proportional to its charge (Z) and inversely proportional to its radius (r). The factor Z/r, the ionic potential, is of great significance not only for the hydration of an ion, but for many of its properties in the presence of water. The importance of ionic potential in sedimentary processes was first pointed out by Goldschmidt as providing a measure of the behavior of an ion toward water. The ionic potentials of a number of ions are given in Table 6.3. In effect, the ionic potential is a measure of electronegativity, since the smaller the radius of a positive ion and the higher its charge, the more acidic is its oxide; and, conversely, the larger the radius and the smaller the charge, the more basic the oxide. From the electrostatic viewpoint, the ionic potential is a measure of the intensity of positive charge on the surface of the ion. This concentration of positive charge on the surface of a cation repels the protons in the coordinated water molecules. If the repulsion is sufficiently great, some of these protons may be detached, thereby neutralizing the charge on the cen-

Table 6.3 Ionic Potentials

Cs^+	-0.60	Th^{4+}	3.9
Rb^+	0.68	Ce^{4+}	4.3
K^+	0.75	Fe^{3+}	4.7
Na^+	1.0	Zr^{4+}	5.1
Li^+	1.5	Be^{2+}	5.7
Ba^{2+}	1.5	Al^{3+}	5.9
Sr^{2+}	1.8	Ti^{4+}	5.9
Ca^{2+}	2.0	Mn^{4+}	6.7
Mn^{2+}	2.5	Nb^{5+}	7.5
La^{3+}	2.6	Si^{4+}	9.5
Fe^{2+}	2.7	Mo^{6+}	9.7
Co^{2+}	2.8	B^{3+}	13
Mg^{2+}	3.0	P^{5+}	14
Y^{3+}	3.3	S^{6+}	20
Lu^{3+}	3.5	C^{4+}	25
Sc^{3+}	3.7	N^{5+}	38

tral cation and resulting in the precipitation of an insoluble hydroxide. With very high repulsive forces, that is, high ionic potentials, all the protons are expelled from the attracted water molecules and an oxyacid anion is formed.

The ionic potential of an element largely determines its place of deposition during the formation of sedimentary rocks and is significant in all mineral-forming processes in an aqueous medium. It provides an explanation for the similar behavior of dissimilar elements as, for example, the tendency of the hydrated ions of bivalent beryllium, trivalent aluminum, and quadrivalent titanium to precipitate together during sedimentation. Elements with low ionic potential, such as sodium, calcium, and magnesium, remain in solution during the processes of weathering and transportation; elements with intermediate ionic potential are precipitated by hydrolysis, their ions being associated with hydroxyl groups from aqueous solutions; elements with still higher ionic potentials form anions containing oxygen which are usually again soluble. When the elements are plotted on a diagram with ionic radius as ordinates and ionic charge as abscissae, the field can thus be divided into three parts: soluble cations, elements of hydrolysates, and elements of soluble complex anions (Figure 6.6).

Some specific examples of the significance of ionic potentials in the behavior of certain elements during sedimentary processes may be cited. Iron in the ferrous state is stable in solution ($Z/r = 2.7$), so that precipitation of iron has to be preceded by oxidation to the ferric state, with a much higher potential ($Z/r = 4.7$). Similarly, manganese is stable in solution as manganous ions ($Z/r = 2.5$) and is precipitated in the hydrated quadrivalent form ($Z/r = 6.7$). Thus in many sedimentary rocks products of hydrolysis and of oxidation are associated in the same deposit. Again, a number of less common and rare quadrivalent elements are concentrated in hydrolysate sediments, their ionic potentials falling within the specified range. For instance, not only beryllium and gallium, which are chemically similar to aluminum, but also titanium, zirconium, and niobium may be concentrated in bauxite, the factor of enrichment being often four- to fivefold as compared to the parent material.
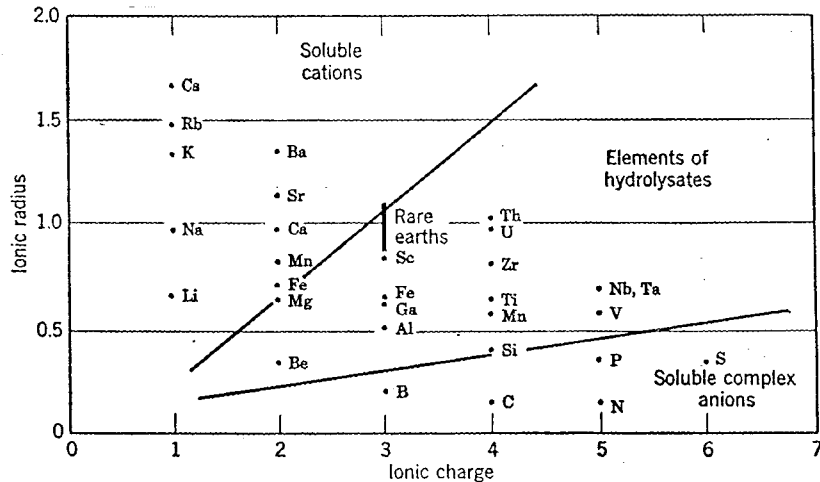
Figure 6.6 Geochemical separation of some important elements on the basis of their ionic potential.

HYDROGEN-ION CONCENTRATION

The hydrogen-ion concentration of natural waters is of great significance in chemical reactions accompanying sedimentary processes. In pure water at 25°C the hydrogen-ion concentration is 10^{-7} mole/liter. If the concentration is greater than that of pure water at the same temperature, the solution is said to be acid; in the contrary case, alkaline. The neutral point alters with rising temperature in the direction of greater hydrogen-ion concentration. As an inverse measure of hydrogen-ion concentration we use the term pH, which is the negative logarithm of this factor; thus the pH of pure water is 7.

The pH of the medium is particularly significant in controlling the precipitation of hydroxides from solution. This is shown in Table 6.4, which gives the pH for beginning precipitation of hydroxides from dilute solutions (about 0.02 M) and the pH of some natural environments. It will be noticed that the sea is slightly alkaline, whereas most terrestrial waters are somewhat on the acid side. The pH of many terrestrial waters is controlled by the buffer system $CaCO_3$-CO_2-H_2O; a saturated solution of CO_2 at its partial pressure of 0.0003 atm in the atmosphere has a pH of 5.7, and a solution of calcite in air-saturated water has a pH near 8.4.

For the transportation and deposition of iron the solubility of ferric hydroxide and the consequent equilibria are of great importance. The solubility product of ferric hydroxide K is given by the equation

$$K = [Fe^{3+}] [OH^-]^3$$

Hence
$$[Fe^{3+}] = \frac{K}{[OH^-]^3}$$

But in water $[OH] = K_w/[H+]$, where K_w is the ionization product of water. Therefore,

$$[Fe^{3+}] = \frac{K \, [H+]^3}{K_w^3}$$

At a fixed temperature K and K_w are constants; hence the concentration of ferric iron in solution is proportional to the cube of the hydrogen-ion concentration. For example, at $18°$ $K = 10^{-38.6}$ and $K_w = 10^{-14.2}$. Therefore, $K/K_{cw}^3 = 10^4$, and at pH $= 7$, $[Fe^{3+}] = 10^{-17}$ mole/liter; at pH $= 6$, $[Fe^{3+}] = 10^{-14}$ mole/liter. In natural waters iron is present not only as Fe^{3+} but also as Fe^{2+} and $FeOH^{2+}$ ions, and the total amount of iron in solution at different pH values is

pH	Amount
8.5	3.10^{-8} mg/m^3
8	4.10^{-7} mg/m^3
7	4.10^{-5} mg/m^3
6	5.10^{-3} mg/m^3

Thus the solubility of iron at pH 6 is about 10^5 times greater than at pH 8.5. Weakly acid iron-bearing solutions flowing into the sea from neighboring land areas must precipitate most of their iron in the weakly alkaline marine waters. This is borne out by actual figures that show that the average content of iron in river waters is about 1 ppm, whereas in sea water the amount is exceedingly small, about 0.008 ppm.

The pH of the environment is especially significant for the transportation of alumina and silica in solution and their ultimate redeposition. To illustrate this, the

Table 6.4 The pH of Natural Media and Its Relation to the Precipitation of Hydroxides

pH	Precipitation of Hydroxides	Natural Media	pH
11			11
	Magnesium		
10		Alkali soils	10
9			9
	Bivalent mangenese		
		Sea water	
8			8
7	Zinc	River water	7
6	Copper	Rain water	6
5	Bivalent iron		5
4	Aluminum	Peat water	4
3		Mine waters	3
	Trivalent iron		
2			2
		Acid thermal springs	
1			
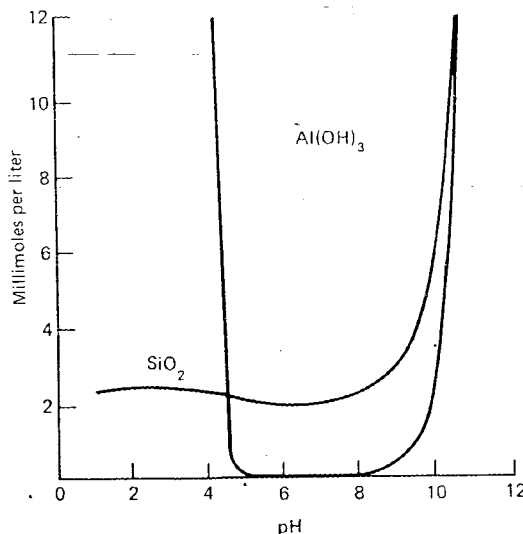
Figure 6.7   The solubility of amorphous silica and
aluminum hydroxide as a function of
pH at 25°C.

solubilities of aluminum hydroxide and silica with respect to pH are plotted on Figure 6.7. At pH < 4 alumina is readily soluble, whereas silica is only slightly soluble. At such pH values alumina would be removed in solution and silica would remain with the parent material. Normal sedimentary environments, however, seldom have a pH this low. From pH 5-9 the solubility of silica increases slightly, but alumina is practically insoluble. Under these conditions removal of silica can take place, leaving alumina behind, as has been inferred during the formation of laterites and bauxites.

An explanation for the formation of either kaolinite or montmorillonite from the same parent materials under different conditions is not clearly apparent from a study of Figure 6.7. The relative solubility of silica and alumina do not change appreciably over a pH range from 4 to 8. It may be that neither mineral is truly an equilibrium phase and each represents an arrested stage in the weathering process. Figure 6.7 would suggest that a true equilibrium assemblage would contain $Al(OH)_3$, a condition found in tropical areas where extensive bauxite deposits may form. Perhaps only under conditions of high rainfall and high temperatures can the weathering reactions tend toward completion before erosion removes the weathered products.

OXIDATION-REDUCTION POTENTIALS

Many elements are present in different oxidation states in the earth's crust. The commonest is iron, occurring as the native metal (oxidation state 0), as ferrous compounds (oxidation state 2), and as ferric compounds (oxidation state 3). Similar elements are manganese (2, 3, 4) sulfur (−2, 0,6), vanadium (3, 4, 5), copper (0, 1, 2), cobalt (2, 3), nitrogen (−3, 0, 5), and many others.

The stability of an element in a particular oxidation state depends on the energy change involved in adding or removing electrons. A quantitative measure for this energy change is provided by the factor known variously as the "oxidation-reduction potential," "oxidation potential," or "redox potential"; we will refer to it as the oxidation potential. The oxidation potential of any reaction is a relative figure, the reference standard being the reaction at 25 °C and 1 atm.

$$H_2 = 2H^+ + 2e \ (e = \text{electron})$$
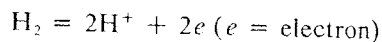
(i.e., the removal of electrons from hydrogen atoms or the oxidation of hydrogen to hydrogen ions). The oxidation potential of this reaction for unit activity (activity is a function of concentration that provides for deviation from the laws of perfect solutions) of the reacting substances is arbitrarily fixed as 0.00 volt, and the scale of oxidation potentials extends on either side of zero. Oxidation potentials are symbolized by E^0 when the relevant reactions take place under the standard conditions of unit activity of the reacting substances, and by Eh when the experimental situation deviates from these conditions.

Table 6.5 lists some reactions in order of decreasing oxidation potential, that is, in order of increasing reducing power, the reduced form of any couple having sufficient energy to reduce the oxidized form of any couple of higher potential. For example, Fe^{2+} reduces Mn^{3+}, and H_2 reduces Fe^{3+}.

The oxidation potential varies with varying concentration of the reacting substances. This variation with concentration is of special importance in reactions involving hydrogen or hydroxyl ions, as many such reactions do (Table 6.5). A variation in pH produces large changes in oxidation potentials involving hydrogen or hydroxyl ions and must be taken into account in applying E^0 values to actual reactions. The

Table 6.5 The Oxidation Potentials of Some Reactions of Geochemical Significance

	E^0 (in volts)
$Co^{2+} = Co^{3+} + e$	1.84
$Ni^{2+} + 2H_2O = NiO_2 + 4H^+ + 2e$	1.75
$Mn^{2+} = Mn^{3+} + e$	1.51
$Pb^2 + 2H_2O = PbO_2 + 4H^+ + 2e$	1.46
$2H_2O = O_2 + 4H^+ + 4e$	1.23
$NH_4^+ + 3H_2O = NO_3^- + 10H^+ + 8e$	0.84
$Fe^{2+} = Fe^{3+} + e$	0.77
$Ni(OH)_2 + 2OH^- = NiO_2 + 2H_2O + 2e$	0.49
$4OH^- = O_2 + 2H_2O + 4e$	0.40
$PbO + 2OH^- = PbO_2 + H_2O + 2e$	0.25
$Co(OH)_2 + OH^- = Co(OH)_3 + e$	0.2
$S^{2-} + 4H_2O = SO_4^{2-} + 8H^+ + 8e$	0.14
$H_2 = 2H^+ + 2e$	0.00
$NH_3 + 9OH^- = NO_3^- + 6H_2O + 8e$	−0.12
$Mn(OH)_2 + OH^- = Mn(OH)^3 + e$	−0.40
$Fe(OH)_2 + OH^- = Fe(OH)_3 + e$	−0.56

A. $H_2 = 2H^+ + 2e$
B. $2H_2O = O_2 + 4H^+ + 4e$
C. $Fe = Fe^{2+} + 2e$
D. $Pb = Pb^{2+} + 2e$
E. $Fe^{2+} = Fe^{3+} + e$
F. $NH_4^+ + 3H_2O = NO_3^-$
$\quad + 10H^+ + 8e$
H. $Pb^{2+} + 2H_2O = PbO_2$
$\quad + 4H^+ + 2e$
K. $Mn^{2+} = Mn^{3+} + e$
L. $Ni^{2+} + 2H_2O = NiO_2$
$\quad + 4H^+ + 2e$
M. $Co^{2+} = Co^{3+} + e$
A'. $H_2 + 2OH^- = 2H_2O + 2e$
B'. $4OH^- = O_2 + 2H_2O + 4e$
C'. $Fe + 2OH^- = Fe(OH)_2$
$\quad + 2e$
D'. $Pb + 2OH^- = PbO$
$\quad + H_2O + 2e$
E'. $Fe(OH)_2 + OH^-$
$\quad = Fe(OH)_3 + e$
F'. $NH_3 + 9OH^- = NO_2^-$
$\quad + 6H_2O + 8e$
H'. $PbO + 2OH^- = PbO_2$
$\quad + H_2O + 2e$
K'. $Mn(OH)_2 + OH^-$
$\quad = Mn(OH)_3 + e$
L'. $Ni(OH)_2 + 2OH^- = NiO_2$
$\quad + 2H_2O + 2e$
M'. $Co(OH)_2 + OH^-$
$\quad = Co(OH)_3 + e$
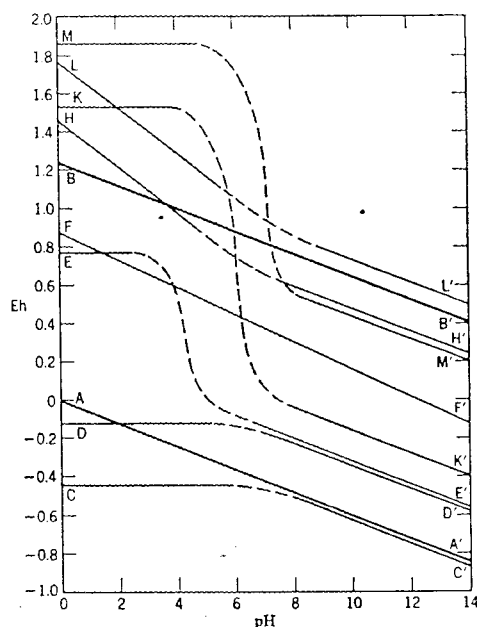
Figure 6.8 Variation of oxidation potential with pH for certain reactions.
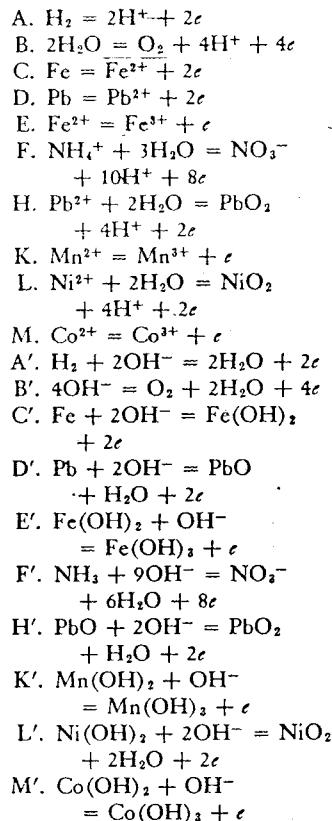
influence of pH on the oxidation potentials of some reactions given in Table 6.5 is shown graphically in Figure 6.8.

The range of oxidation potentials of natural environments determines the reactions that may take place. Chemical reactions in aqueous media are theoretically limited to those with oxidation potentials between those for the reactions

$$H_2O = \tfrac{1}{2}O_2 + 2H^+ + 2e, \quad E^0 = 1.23 \text{ volts} \tag{1}$$

$$2H^+ + 2e = H_2, \quad E^0 = 0.00 \text{ volt} \tag{2}$$
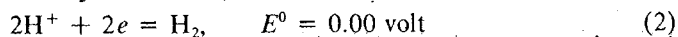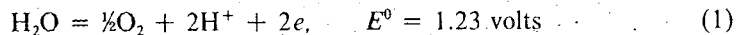
The oxidized form of any couple with a higher potential than that for (1) will theoretically decompose water with the evolution of oxygen. The reduced form of any couple with a lower potential than that for (2) will theoretically decompose water with the evolution of hydrogen. These requirements are not strictly met in practice on account of overvoltage phenomena; that is, it requires a greater potential than the theoretical to produce the evolution of hydrogen or oxygen at a measurable rate. However, the chemistry of sedimentation indicates that the potentials of these two reactions do largely control oxidation and reduction under natural conditions.

Both these reactions involve hydrogen ions, and their potentials are thus strongly affected by changes in pH. The E^0 values given above are for hydrogen-ion concentrations of unity, that is, pH = 0, and the potentials decrease (at 25°) 0.06 volt for each unit increase in pH. The pH of natural waters is variable, ranging from as low as 0 in strongly acid waters of volcanic regions to 10 or more in alkaline areas where sodium carbonate is present in solution. As a rule, however, the pH of natural waters lies between 4 and 9, the great majority being within one unit of the figure for pure water (pH = 7). For a pH of 7 the potential of (1) is 0.82 volt and of (2) — 0.41 volt. [The potential of 0.82 volt is that of water with pH 7 saturated with oxygen at 760 mm pressure; for oxygen at its partial pressure in the atmosphere (160 mm) this potential is reduced to 0.81 volt.] These figures indicate that the oxidation potentials of natural environments, where the pH is near 7, should lie between — 0.41 and 0.82 volt. Measurements of oxidation potentials in natural waters are in agreement with this statement (Figure 6.9) although values as low as — 0.5 volt have been recorded
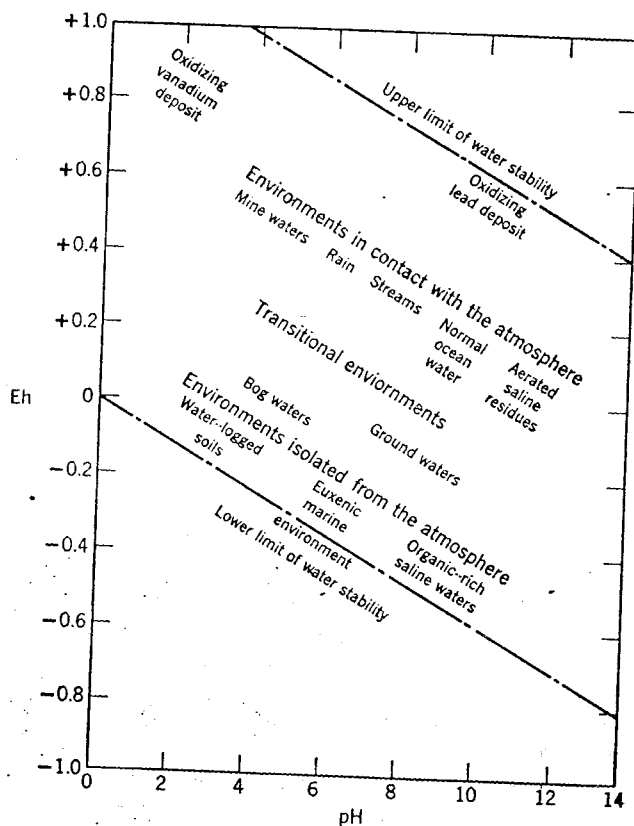
Figure 6.9 Approximate position of some natural environments as characterized by Eh and pH. (From Garrels and Christ, *Solutions, Minerals, and Equilibria.* Courtesy Harper and Row)

in marine-bottom deposits rich in organic matter; hydrogen may actually be generated in such environments.

The solution, transportation, and deposition of elements that may occur in two or more oxidation states are directly and powerfully influenced by the oxidation potential of the environment. For example, the high potentials required to convert bivalent to trivalent cobalt and bivalent lead to PbO_2 in acid solution indicate that the minerals stainierite ($CoOOH$) and plattnerite (PbO_2) are deposited from alkaline solutions, for which the oxidation potentials are much less (Figure 6.8). Figure 6.8 shows that for many reactions the oxidation potentials decrease rapidly with the increase of pH; generally, therefore, oxidation proceeds more readily the more alkaline the solution. The potentials for some oxidations in alkaline solutions, particularly those that result in the precipitation of almost insoluble compounds, lie far below the potentials for corresponding oxidations in acid solutions. This is especially marked with respect to the oxidation of ferrous to ferric iron; in acid solution the potential is 0.77 volt and is not affected by pH; however, as soon as the pH increases to a figure at which ferric hydroxide is precipitated, the oxidation potential drops sharply to a negative figure. Thus ferrous salts are comparatively stable in acid solution, being only slowly oxidized by air, but in solutions sufficiently low in acid for Fe $(OH)_3$ to be precipitated oxidation proceeds rapidly to completion. Deposition of ferrous compounds in nature therefore demands either a very acid environment or one with a very low oxidation potential, on the negative side of zero.

The Eh—pH diagrams provide a useful device for illustrating the stability fields of different minerals in an aqueous environment. These diagrams are discussed in detail by Garrels and Christ (1965), with numerous examples illustrated by them. The system Mn—H_2O under varying conditions of Eh and pH provides an illustrative case (Figure 6.10), the diagram showing clearly why manganese has not been found as the native metal. Under reducing conditions in the geological environment manganese occurs as manganous compounds in acid solutions, and precipitates as Mn $(OH)_2$ when the pH exceeds 8 (for unit activity of Mn; for more dilute solutions, as found in nature, the pH required for precipitation is greater, according to the law of mass action). As conditions become more oxidizing, the field of manganous ions shrinks toward more acid conditions; on the alkaline side of the diagram Mn $(OH)_2$ (pyrochroite) is successively replaced by Mn_2O_4 (hausmannite), MnOOH (manganite), and MnO_2 (pyrolusite). Under highly oxidizing conditions MnO_2 is the stable phase over the whole range of pH. The diagram also explains why permanganates do not occur under geological conditions; aqueous solutions of permanganate are unstable and slowly decompose, liberating oxygen and precipitating MnO_2.

The separation of closely related elements in the upper zone of the lithosphere by processes involving solution and redeposition is often brought about by their distinctive properties with respect to oxidation and reduction. Thus the three elements iron, nickel, and cobalt often occur together in primary deposits, yet supergene, or near-surface aqueous processes result in their separation. These three elements differ greatly in the potentials required to oxidize them beyond the bivalent state. Iron is readily oxidized to the trivalent state in alkaline and mildly acid environments; cobalt requires a much higher potential even in alkaline solution, and in acid solutions the
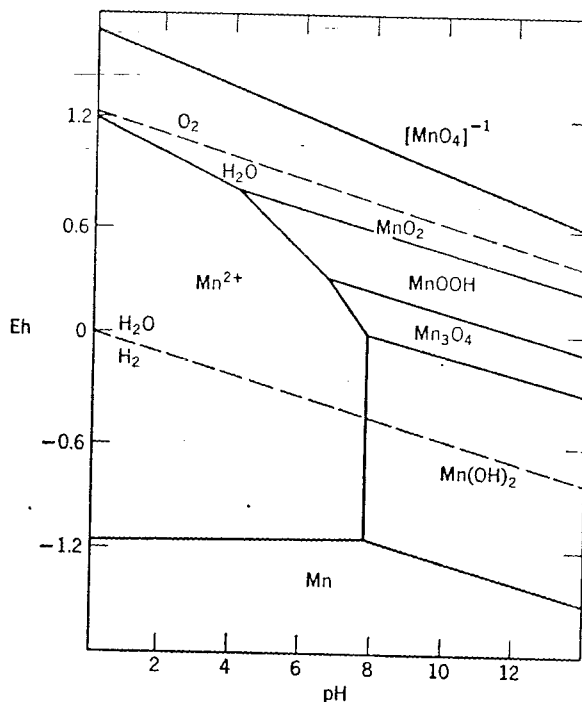
Figure 6.10   Eh-pH diagram for the system Mn-H_2O, for unit activity of Mn. The dashed lines are for the reactions for the decomposition of water (AA^1 and BB^1 of Figure 6.7).

potential required lies high above that for the release of oxygen from water; nickel does not form trivalent compounds, but a dioxide is known, the formation of which even in alkaline solutions requires potentials somewhat higher than that for the release of oxygen from water. This is reflected in natural occurrences: the common form of iron in supergene deposits is hydrated ferric oxide; hydrated cobaltic oxide (stainierite) is found only where conditions have been strongly oxidizing; and the higher oxide of nickel is not known as a mineral. The separation of these three elements by supergene processes is well illustrated where intense weathering of ultrabasic rocks has given rise to lateritic material rich in Fe_2O_3, concentration of the nickel as garnierite, and of the cobalt as hydrated cobaltic oxide or as cobaltian wad.

Oxidation processes also result in a similar separation of manganese from iron. Manganese is often present in solid solution in primary minerals containing iron, the ferrous and the manganous ions being mutually replaceable. Supergene processes, however, generally lead to fairly complete separation of iron from manganese, since the potential required to convert iron to the ferric state is much lower than that required to convert manganese to manganese dioxide. The iron readily precipitates as hydrated ferric oxide, whereas the manganese remains in solution longer and is
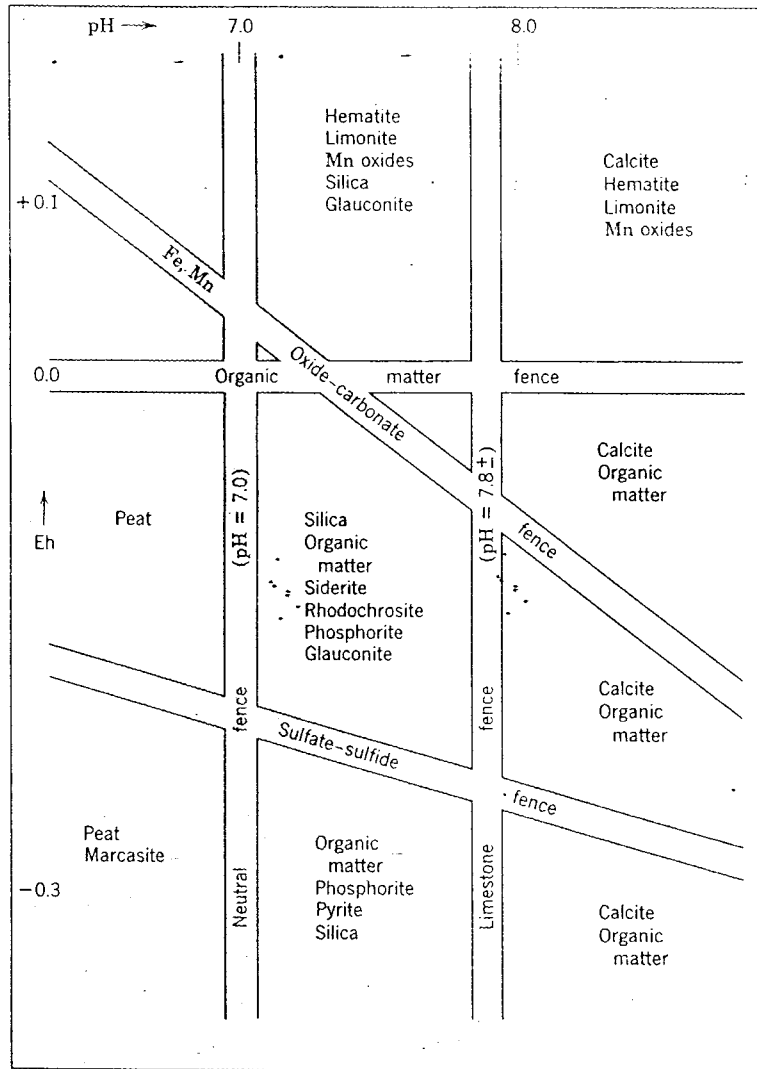
Figure 6.11 Sedimentary associations in relation to environmental limita-
tions imposed by oxidation potential and pH. (After Krumbein
and Garrels, *J. Geol.* **60**, 26, 1952)

eventually deposited under more oxidizing conditions as comparatively iron-free
manganese dioxide.

Oxidation potentials and pH are the basic controls that determine the nature of
many sedimentary products. Krumbein and Garrels (1952) have devised an inge-
nious diagram which illustrates the relation between these factors and the geological
materials on which they act (Figure 6.11). In this diagram they develop the concept
of the "geochemical fence," a boundary defined by the presence of a particular min-

eral or material on one side and its absence on the other, in effect, by a certain chemical reaction. A particular geochemical fence may represent a specific pH value or a specific oxidation potential or a combination of both factors. Many geochemical fences can be defined, but Krumbein and Garrels have shown that the most generally useful ones in considering sedimentary processes are the neutral fence, at pH = 7; the limestone fence, at pH = 7.8 (at higher pH calcite is readily deposited, at lower pH it tends to dissolve); the sulfate-sulfide fence, determined by the sulfide-sulfate oxidation potential; the Fe, Mn oxide-carbonate fence, determined by the oxidation potential at which ferrous and manganous compounds (mainly carbonates in the sedimentary environment) oxidize to the higher oxides; and the organic matter fence, below which organic matter is stable and above which it oxidizes to carbon dioxide. As a result, we obtain a classification of sedimentary environments based on the two significant parameters of pH and oxidation potential, and even though the bulk chemistry of sedimentary deposits may vary greatly, the careful study of their mineralogy will elucidate the physicochemical conditions under which they developed.

COLLOIDS AND COLLOIDAL PROCESSES

The colloidal state is one of fine subdivision, the size range being approximately 10^{-3} to 10^{-6} mm. Colloidal solutions grade into true solutions on the one hand and into suspensions on the other without any distinct line of demarcation. The degree of dispersion is usually greater than the resolving power of an ordinary microscope but is less than molecular; that is, colloidal particles are generally multimolecular. The colloid particles of the *disperse* phase are separated by the *dispersion* medium; the whole may be referred to as a disperse system. There are a number of different types of disperse or colloidal systems [solid-gas (smokes), liquid-gas (fogs), liquid-liquid (emulsions)], but the important type in sedimentary processes is a solid-liquid system (sols, gels, and pastes) in which the liquid is water. Sols are systems that resemble liquids in their physical properties; thus they flow readily and do not show rigidity. Gels, on the other hand, show some ridizity. Pastes are systems in which, although the solid part is in the form of discrete particles, the concentration of these particles has been so much increased that they form the bulk of the system, for example, plastic clay.

Sols in water are divided into two types, hydrophilic and hydrophobic. In the first type there is strong interaction between the particles and the water molecules, which serves to stabilize the colloidal solution. In the second type there is no such attraction between the particles and the water molecules, and they are thus much less stable than hydrophilic sols and more easily precipitated. In general the particles in hydrophobic sols are larger than those in hydrophilic sols. Further, when a hydrophobic sol has been precipitated by some physical change it cannot be readily reconverted into a sol by reversing this change, whereas a hydrophilic sol is generally reversible in this respect. No sharp boundary exists between these two types of sols, but the division is useful. Silica is an example of a substance that forms a hydrophilic sol; aluminum hydroxide, on the other hand, forms a hydrophobic sol.

Colloidal particles are electrically charged. This charge may originate in two ways:

either by adsorption of ions from the liquid or by the direct ionization of the material of the particle. Some colloids, for example, ferric hydroxide, may acquire either positive or negative charges according to the environment in which they are formed. The charges on some important colloids are as follows:

Positive	Negative
Aluminum hydroxide	Silica
Ferric hydroxide	Ferric hydroxide
Chromic hydroxide	Vanadium pentoxide hydrate
Titanium dioxide hydrate	Manganese dioxide hydrate
Thorium dioxide hydrate	Humus colloids
Zirconium dioxide hydrate	Sulfide sols

Colloidal particles may be either crystalline or amorphous. Most inorganic colloids are crystalline. The establishment of this fact has been an important contribution of X-ray diffraction techniques to the study of colloids.

Colloids may be produced by two procedures; either coarser particles may be broken down to colloidal dimensions, or smaller particles (molecules or ions) may be aggregated into particles of colloidal size. Most naturally formed colloids probably originated in the second way.

The colloidal state is always a metastable one. An increase in the size of the particles will lead to a decrease in the total surface area, and hence a decrease in the free energy of the system, so that all colloids are theoretically unstable, although they may remain unchanged over long periods. The charge on the particles is an important factor in the stability of a colloid, and this generally requires the presence of small concentrations of electrolytes. Large amounts of electrolytes, however, are precipitants of colloids; on this account most colloidal matter is rapidly flocculated by sea water. Colloids show great differences in relative stability. Some are stable under wide variations in chemical and physical environments; others require carefully controlled conditions in order to exist in the colloidal state at all. Obviously only those substances that form rather stable colloids are of importance in geological processes. Unstable colloids can, however, be stabilized by the presence of other substances. Of these the most important are organic compounds. In the geological environment such compounds are generally referred to as humus colloids, for lack of more precise information as to their nature; they are probably albumins. The presence of these humus colloids appears to play an important part in stabilizing the inorganic colloids formed during sedimentary processes, thus enabling transport of such material over much greater distances than would otherwise be possible.

Because silica is the most abundant material in the earth's crust and it readily goes into colloidal solution, the role of colloidal silica in sedimentary processes has been the subject of a great deal of discussion. Much of the geological literature relating to the transportation and precipitation of dissolved silica has been essentially unanimous as to the colloidal state of the silica. However, the development of colorimetric tests that permit discrimination between colloidal silica and silica in ionic solution has shown that this conclusion is based on inadequate evidence. These tests have shown that most of the silica in natural waters is in true solution. The significance of this fact for the dissolution and precipitation of silica during sedimentation

has been carefully discussed by Krauskopf (1956). He shows that the origin of chert may be plausibly ascribed to dissolution of remains of siliceous organisms and to reprecipitation of the silica (initially in an amorphous state), but not in general to direct inorganic precipitation.

An important property of colloidal particles is their ability to bind and concentrate certain substances through adsorption. Two types of adsorption are recognized: (a) physical or van der Waals' adsorption, and (b) chemical adsorption. Both types may act together, and all gradations between extremes exist. Physical adsorption is characterized by low heats of adsorption and by a loose bonding of the adsorbate to the adsorbent. Chemical adsorption, or chemisorption, on the other hand, is characterized by high heats of adsorption and a firm chemical bonding (i.e., by valence bonds) of the adsorbate. It may involve the bonding of a foreign cation or a foreign anion, or both, to open bonds at the surface of the adsorbent; or it may be the exchange or substitution of a foreign cation or anion, or both, for a cation or anion at the surface. The property of base exchange, shown particularly by clays, whereby cations in the clay may be exchanged for other cations present in aqueous solutions in contact with the clay minerals, may be considered as a particular type of adsorption. Some principles governing adsorption may be stated as follows:

1. The amount of adsorption increases as the grain size of the adsorbent decreases, and hence its surface area increases.
2. Adsorption is favored if the adsorbate forms a compound of low solubility with the adsorbent (an example is the adsorption of phosphate ions by ferric hydroxide).
3. The amount of a substance adsorbed from solution increases with its concentration in that solution.
4. Highly charged ions are adsorbed more readily than lower charged ions.

Through adsorption processes many ions may be removed from natural waters. The clay minerals, especially montmorillonite, show a marked adsorptive capacity; the chemical adsorption of potassium ions by montmorillonite may result in the formation of illite. Many complex ions (for example, those containing arsenic and molybdenum) and ions of the heavy metals are also adsorbed and thus removed from solution by natural colloids. Such elements are often enriched in sedimentary iron and manganese ores. Here we are dealing with a systematic "depoisoning" of the hydrosphere (analogous to the "depoisoning" effect of freshly precipitated ferric hydroxide, utilized in medicine); without this phenomenon a number of biologically damaging elements would accumulate in ocean waters. Sufficient amounts of many elements, such as copper, selenium, arsenic, and lead, have potentially been supplied by weathering and erosion during geological time to cause serious poisoning of the ocean had not some process of elimination of these substances been active.

PRODUCTS OF SEDIMENTATION

A cursory examination of sedimentary processes suggests that they would tend to produce an average mixture of the individual components present in the parent material and thus work against any chemical differentiation. This, however, is not

the case; weathering, erosion, and sedimentation lead generally to a marked separation of the major elements. As Goldschmidt pointed out, the cycle of matter at the earth's surface can be likened to a chemical analysis and to a quantitative analysis at that. The chemical differentiation that results is remarkable. The steps in this geochemical separation process are as follows:

1. Minerals that are especially resistant to chemical and mechanical breakdown collect as granular material. Of these, the commonest is quartz, and the product is a quartz sand or a sandstone showing a marked enrichment in silicon with respect to the parent material. This may be compared to the separation of silica in the first stage of a rock analysis.

2. Accumulation of the products of chemical breakdown of aluminosilicates, giving a mud consisting essentially of the clay minerals. This results in concentration of aluminum and also of potassium by adsorption. The process corresponds to the second step in a rock analysis, the separation of alumina and other easily hydrolyzed bases.

3. Along with the formation of argillaceous sediments, but often separated in space and time, iron is precipitated as ferric hydroxide. In this process oxidation from the ferrous to the ferric state precedes precipitation by hydrolysis. Concentration of iron is the result, sometimes to the extent of the formation of iron ores.

4. Calcium is precipitated as calcium carbonate either by purely inorganic processes or by the action of organisms. Limestones are formed and calcium thereby concentrated. This may result in almost quantitative separation of calcium, as in a chemical analysis. Limestone can be partly or wholly converted to dolomite by the metasomatic action of magnesium-rich solutions and magnesium thereby precipitated and concentrated together with calcium.

5. The bases that remain in solution collect in the ocean, from which they are removed in quantity only by evaporation, giving rise to salt deposits. The most important of these bases is of course sodium, but lesser amounts of potassium and magnesium also accumulate in sea water.

The chemical breakdown of a rock by weathering can be represented by the following scheme:

Si	Al, Si, (K)	Fe	Ca, (Mg)	(Ca,) Na, (K), (Mg)
Resist-ates SiO_2	Hydroly-sates, clay minerals	Oxidates $Fe(OH)_3$	Carbonates $CaCO_3$ $CaMg(CO_3)_2$	Evaporates NaCl $CaSO_4$ $MgSO_4$ etc.
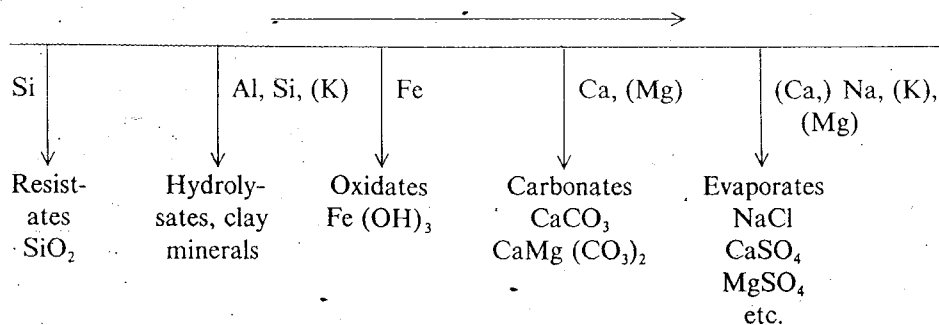
This scheme indicates the course followed by the major elements during sedimentation and gives a useful geochemical classification of sediments into resistates,

hydrolysates, oxidates, carbonates, and evaporates. Goldschmidt recognized a further class, the reduzates, which includes coal, oil, sedimentary sulfides, and sedimentary sulfur. Coal and oil are of organic origin and are discussed in Chapter 10. Sedimentary sulfides and sulfur also often owe their formation to organic processes, although possibly indirectly.

The fate of the major elements during sedimentation has been fairly well worked out. Silica concentrates in the resistate sediments, alumina in the hydrolysates, iron and manganese in the oxidates, and calcium and magnesium in carbonates. A good part of the sodium remains in solution and eventually accumulates in the ocean; potassium is adsorbed by the clays and may form the minerals illite and glauconite. The fate of the minor elements during sedimentation has not been thoroughly investigated, and much less is known about their behavior under these circumstances than during magmatism. Sedimentary processes are more complex and less uniform than igneous processes and many factors may play a part in determining the the transportation and deposition of an element; they include ionic potential, pH, and oxidation potential of the medium, collodial properties, and adsorption. Hence it is not yet possible to make categorical statements regarding the fate of many of the minor elements during sedimentation.

Turekian and Wedepohl have prepared a critical compilation of the available data for the three principal groups of sedimentary rocks—shales, sandstones, and carbonates. Their figures, updated where necessary, are given in Table 6.6 together with the averages for igneous rocks for comparison. The data indicate that the common sedimentary rocks seldom show marked enrichment in minor and trace elements over the amounts present in igneous rocks—in fact, sandstones and carbonates are usually depleted in these elements. Most trace elements are somewhat more concentrated in shales than in other types of sedimentary rocks, but the amount in ordinary shales seldom exceeds the igneous rock average.

The resistates form the important group of sands and sandstones. Quartz is by far the commonest and most abundant of residual minerals, and sands and sandstones are sources of silica for industrial uses. Many other minerals may appear in small amounts in sands and sandstones, but most of these can be decomposed and removed by intense weathering. Zircon is one of the most persistent of minerals, and the main ore deposits of zircon are sands from which it can be profitably separated. Magnetite and ilmenite are fairly resistant minerals and accumulate in sands; other industrially important constituents of some sands are rutile, monazite, cassiterite, and, of course, gold and the platinum metals.

The hydrolysate sediments consist in great part of the clay minerals. Tropical weathering often produces aluminum hydroxides rather than hydrated aluminum silicates, and high-alumina clays and bauxites result. In either process the end product represents a concentration of aluminum over the average amount in the earth's crust. As was indicated in the discussion of ionic potential, many elements, especially those in groups III and IV of the periodic table, may be expected to precipitate in hydrolysate sediments, and the meager data on minor elements in sedimentary rocks bear this out. Shales show concentrations of elements of medium ionic potential and of elements like potassium that are readily adsorbed by colloidal particles; they are sometimes enriched in chalcophile elements, probably precipitated as sulfides by the

Table 6.6   Abundances of the Elements (in Parts Per Million) in the Principal Types of
Sedimentary Rocks

Element	Shales	Sandstones	Carbonates	Igneous Rocks
Li	66	15	5	20
Be	3	0.X	0.X	2.8
B	100	35	20	10
F	740	270	330	625
Na	9,600	3,300	400	28,300
Mg	15,000	7,000	47,000	20,900
Al	80,000	25,000	4,200	81,300
Si	273,000	368,000	24,000	277,200
P	700	170	400	1,050
S	2,400	240	1,200	260
Cl	180	10	150	130
K	26,600	10,700	2,700	25,900
Ca	22,100	39,100	302,300	36,300
Sc	13	1	1	22
Ti	4,600	1,500	400	4,400
V	130	20	20	135
Cr	90	35	11	100
Mn	850	X0	1,100	950
Fe	47,200	9,800	3,800	50,000
Co	19	0.3	0.1	.25
Ni	68	2	20	75
Cu	45	X	4	55
Zn	95	16	20	70
Ga	19	12	4	15
Ge	1.6	0.8	0.2	1.5
As	13	1	1	1.8
Se	0.6	0.05	0.08	0.05
Br	4	1	6.2	2.5
Rb	140	60	3	90
Sr	300	20	610	375
Y	26	15	6.4	33
Zr	160	220	19	165
Nb	11	0.0X	0.3	20
Mo	2.6	0.2	0.4	1.5
Ag	0.07	0.0X	0.0X	0.07
Cd	0.3	0.0X	0.09	0.08
In	0.1	0.0X	0.0X	0.1
Sn	6.0	0.X	0.X	2
Sb	1.5	0.0X	0.2	0.2
I	2.2	1.7	1.2	0.5
Cs	5	0.X	0.X	3
Ba	580	X0	10	425
La	24	16	6.3	30
Cc	50	30	10	60
Pr	6.1	4.0	1.5	8.2
Nd	24	15	6.2	28
Sm	5.8	3.7	1.4	6.0
Eu	1.1	0.8	0.3	1.2
Gd	5.2	3.2	1.4	5.4

Table 6.6   (Continued)

Element	Shales	Sandstones	Carbonates	Igneous Rocks
Tb	0.9	0.6	0.2	0.9
Dy	4.3	2.6	1.1	3.0
Ho	1.2	1.0	0.3	1.2
Er	2.7	1.6	0.7	2.8
Tm	0.5	0.3	0.1	0.5
Yb	2.2	1.2	0.7	3.4
Lu	0.6	0.4	0.2	0.5
Hf	2.8	3.9	0.3	3
Ta	0.8	0.0X	0.0X	2
W	1.8	1.6	0.6	1.5
Hg	0.4	0.3	0.2	0.15
Tl	1.0	0.5	0.2	0.8
Bi	0.4	0.17	0.2	0.1
Pb	20	7	9	13
Th	12	1.7	1.7	9.6
U	3.7	0.45	2.2	2.7

Turekian and Wedepohl, *Bull. Geol. Soc. Am.*, **72**, 175, 1961, with more recent data. The figures for igneous rocks are given for comparison. (For some elements only order of magnitude estimates are made; these are indicated by the symbol X.

H_2S often generated in marine muds. The most remarkable hydrolysate sediments from the geochemical viewpoint are the black bituminous shales and the bauxites, both of which have originated under rather special conditions. The black shales were deposited slowly in a strongly reducing marine environment rich in organic matter. Sulfide ions were evidently present, produced by the reduction of sulfate. Analyses of black shale are characterized not only by a considerable content of organic carbon, but generally also by much sulfur, present mainly as FeS_2. Enrichment of the following minor elements has been noted: V, U, As, Sb, Mo, Cu, Ni, Cd, Ag, Au, and metals of the platinum group. Vanadium has been produced commercially from such shales, and they are far more significant than all the primary deposits as a potential source of large amounts of uranium. It has been suggested that the minor elements were accumulated by the vital activity of the organisms now represented by bituminous material, but this is far from certain. Judging from the chalcophile nature of many of these elements, precipitation from solution as sulfides seems a more reasonable explanation. A linear increase of uranium with increasing carbon content has been demonstrated in some of these shales, but this does not necessarily imply that the uranium was present in the organisms that furnished the carbon. The uranium content also shows an excellent correlation with the abundance of colloidal size grades in the sediment; this might suggest that the uranium is present in the clay mineral, which in these black shales is generally illite. The evidence indicates that the concentration of uranium is not the result of biological activity but of later chemical processes probably related, in part at least, to the presence of organic matter in the sediments. Some phosphatic shales, such as those of the Phosphoria formation in

Wyoming, Idaho, and Montana, show similar geochemical features, especially in the enrichment in vanadium and uranium; they seem to have been deposited under similar conditions, that is, oxygen-deficient marine environments where organic material was accumulating and the rate of sedimentation was very slow. Bauxites have a different pattern of enrichment and often show a concentration of beryllium, gallium, niobium, and titanium; of these gallium is actually being extracted as a byproduct from the production of aluminum.

The most important oxidate is ferric hydroxide, which, if pure, gives rise to a sedimentary iron ore. Manganese is also deposited as an oxidate sediment in the form of hydrated manganese dioxide, and such deposits or their metamorphosed equivalents are the significant sources of manganese ore. The adsorptive power of precipitated ferric hydroxide and manganese dioxide hydrate is very great; hence many minor elements are found in oxidate sediments. The pattern of enrichment differs; hydrated manganese dioxide, being a negatively charged colloid, adsorbs cations, whereas ferric hydroxide, generally a positively charged colloid, adsorbs anions. These adsorbed elements are sometimes in sufficient amounts to be important commercially, either as profitable byproducts (e.g., Ni in some sedimentary iron ores, W in some manganese ores) or as deleterious impurities (e.g., As in some iron ores). Vanadium, phosphorus, arsenic, antimony, and selenium have been reported in sedimentary iron ores in larger amounts than their average abundance in the crust; Li, K, Ba, B, Ti, Co, Ni, Cu, Mo, As, V, Zn, Pb, and W have been reported in notable concentrations in manganese ore.

The common carbonate sediment is, of course, limestone, which consists essentially of calcite. Calcium carbonate may also be deposited as aragonite, but it is doubtful whether aragonite will persist for any considerable time in a geological formation, since it tends to change to calcite. Whether calcium carbonate was originally deposited as calcite or as aragonite may have significant geochemical consequences; the structure of aragonite permits ready substitution by larger cations, such as strontium and lead, but not the smaller cations, whereas for calcite the reverse is true. Hence the minor elements in a limestone will probably differ in kind and amount according to the nature of the calcium carbonate in·the original sediment.

The evaporates are quantitatively unimportant as sediments but are highly significant in the interpretation of geological history. Geochemically they are of special interest as a type of deposit with a mode of formation that can readily be reproduced in the laboratory. They have been well described by Stewart (1963). As early as 1849 Usiglio made experiments aimed at elucidating the conditions of formation of salt deposits, but his results were unsatisfactory, because he worked with sea water, a highly complex solution with which he failed to get reproducible results. Later the problem was tackled from the other direction by van't Hoff and his co-workers, who began by studying the solubility relations of all the possible compounds that might be produced by the evaporation of sea water. Working initially at $25°$, they determined the equilibrium relations in the simple two-component salt-water systems and then extended these researches to multicomponent systems. Similar investigations were made at $83°$, and specific reactions involving the appearance or disappearance

of individual compounds were studied at the temperature of reaction. Van't Hoff's success in working out phase relations and applying these results to natural occurrences of evaporates (especially the Stassfurt deposits) was one of the first fruits of the application of physicochemical principles (in this case the phase rule) to geological problems.

As sea water evaporates under natural conditions, calcium carbonate is the first solid to separate. The precipitation of calcium carbonate may be followed by that of dolomite, but there is no evidence that extensive deposits of dolomite have been formed in this way. Indeed, evaporation of sea water in a closed basin cannot give rise to thick carbonate deposits—sea water 1000 m deep would give only a few centimeters of limestone.

With continued evaporation calcium sulfate is deposited. Depending on temperature and salinity, either gypsum or anhydrite may be formed. In salt solutions of approximately the composition of sea water at 30° gypsum will begin to separate when the salinity has increased to 3.35 times the normal value; after nearly one-half the total amount of calcium sulfate has been deposited anhydrite becomes the stable phase. When the solution has been concentrated to one-tenth of the original bulk, halite starts to separate. Anhydrite and halite then precipitate together until the field of stability of polyhalite, $K_2Ca_2Mg(SO_4)_4 \cdot 2H_2O$, is reached.

Most evaporate deposits contain calcium carbonate, calcium sulfate, and sodium chloride; evidently conditions under which other salts could be deposited have seldom been attained. Only when an evaporating body of sea water has been reduced to 1.54% of the original volume do potassium and magnesium salts begin to crystallize. Important deposits of these salts are worked in Germany, in the Texas–New Mexico area of the United States, in the province of Perm in the U.S.S.R., and in Saskatchewan, Canada.

The further course of crystallization can be discussed in terms of a triangular diagram having corners which represent K_2, Mg, and SO_4 and which shows the compounds in equilibrium with halite and saturated solutions at a particular temperature. Figure 6.12 is such a diagram for 25°. It shows that the continued isothermal evaporation of sea water leads to the crystallization of blödite, $Na_2Mg(SO_4)_2 \cdot 4H_2O$. The composition of the solution follows a path directly away from the point corresponding to blödite and soon crosses the boundary into the field of epsomite, $MgSO_4 \cdot 7H_2O$. Epsomite then begins to crystallize, and the further course depends on whether the previously separated blödite can react with the solution. If it does, it will be resorbed; under natural conditions, however, it may be crusted over and effectively removed from the system. In either event, the path of crystallization eventually reaches the boundary between the fields of epsomite and kainite, $KMgSO_4Cl \cdot 3H_2O$, and then follows this boundary curve, the two salts crystallizing together. The succession is hexahydrite-kainite, kieserite-kainite, kieserite-carnallite, and finally the three salts kieserite-carnallite-bischofite separate until evaporation is complete. During all this, halite and small amounts of calcium salts are still being deposited. Polyhalite, which began to crystallize almost contemporaneously with halite, ceases to form during the course of separation of kainite.

The accompanying table illustrates the theoretical profile to be expected from the evaporation of sea water at 25°. This theoretical profile shows a general correspondence with the succession in natural salt deposits in most areas. Commonly one finds limestones and dolomites passing up into calcium sulfate and halite, with or without polyhalite. The usual form of calcium sulfate in natural deposits is anhydrite rather

Kieserite, carnallite, bischofite			Bischofite zone
Kieserite, carnallite	Anhydrite		Carnallite zone
Kieserite, kainite			
Hexahydrite, kainite			Kainite zone
Epsomite, kainite		Halite	
Epsomite			Potash-free magnesium sulfate
Blödite	Polyhalite		zone
Polyhalite			Polyhalite zone
Anhydrite			Anhydrite zone
Gypsum			Gypsum zone
Carbonates			Basal limestone and dolomite

than gypsum. The correspondence with the higher parts of the theoretical profile is less marked, however. For example, sylvite, which does not appear in this profile, is the important potassium salt in many deposits. Epsomite and hexahydrite have generally lost water of crystallization and have been converted into kieserite. A bischofite zone is rarely developed, probably because complete evaporation is seldom if ever achieved in nature. Van't Hoff's investigations at higher temperatures showed that most of the mineralogical features of natural deposits can be explained by assuming either a temperature of evaporation considerably higher than 25° or recrystallization at some higher temperatures after burial—a mild thermal metamorphism. The latter suggestion is highly plausible.

TOTAL AMOUNT OF SEDIMENTATION

The total amount of sedimentation during geological time is clearly a figure of great importance for quantitative geochemistry, and a number of attempts have been made to calculate it. Such calculations are generally based on data regarding the amount and composition of ocean water and the average composition of igneous and sedimentary rocks. If we can assume that the total sodium content of sea water is derived from the weathering and erosion of igneous rocks, this quantity will be proportional to the total amount of sediments. Clarke used the following data:

Sodium content of the ocean	1.14%
Average sodium content of igneous rocks	2.83%
Composition of the 10-mile crust	93% lithosphere, 7% hydrosphere
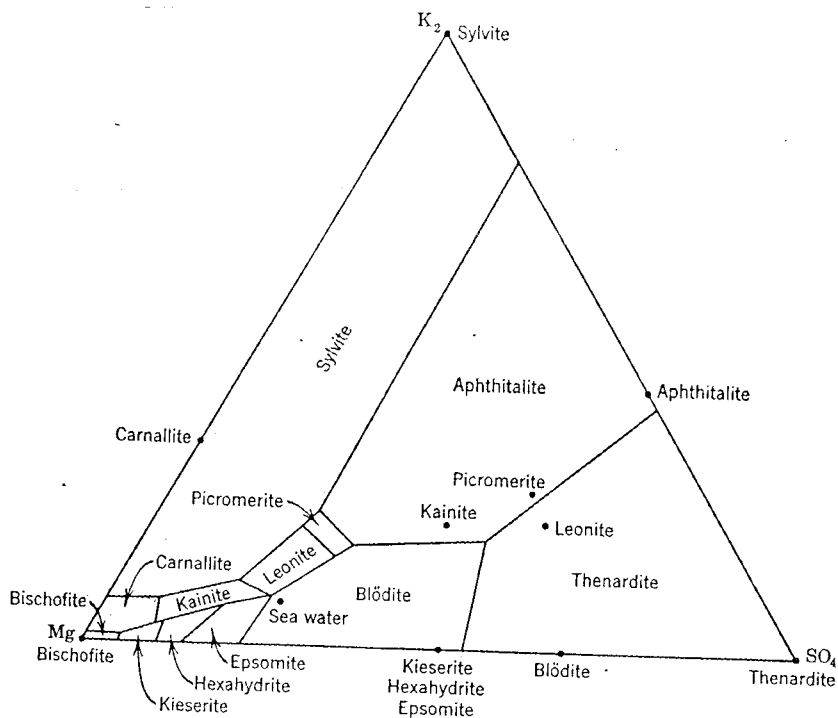
Figure 6.12 Stability fields of evaporate minerals at 25°. (After Stewart, 1963)

From these figures the ratio between oceanic sodium and rock sodium is about 1:30. Hence the sodium in the ocean is equivalent to that contained in 1/30 of all the igneous rocks of the earth's crust to a depth of 10 miles. This fraction amounts to 54.8×10^6 cubic miles. However, the calculation is subject to correction, for it rests on the assumption that all the sodium accumulates in the sea. This is not so; the sedimentary rocks contain sodium, even if in much smaller amount than the parent igneous rocks. Clarke put the average sodium content of sedimentary rocks at 0.90%, or 35% of the average in igneous rocks. Thus, of the original sodium, less than 65% has come to rest in the sea. The estimate of the amount of igneous rocks required to give the present sodium content of the ocean must therefore be increased by a factor of at least 100/65, from 54.8×10^6 cubic miles to 84.3×10^6 cubic miles. An allowance for a 10% increase of volume of sedimentary rocks from the parent igneous material by oxidation, carbonation, and hydration gives a figure of about 93×10^6 cubic miles, or 3.7×10^8 cubic kilometers for the total volume of sediments produced during geological time. This corresponds to a rock shell nearly 2500 feet thick enveloping the whole earth. If the material were confined to the continental platforms (roughly one-third of the area of the globe), its thickness would be about 7300 feet.

Goldschmidt used a somewhat different procedure to calculate the total amount of igneous rock weathered during geological time and the amounts of the different sediments. For each square centimeter of the earth's surface, there are 278 kg of sea

water; and because sea water contains 1.07% sodium (a later figure than Clarke's 1.14%), the 278 kg contains 2.975 kg of sodium. The average sodium content of igneous rocks is 2.83% and of sedimentary deposits approximately 1%. In the process of weathering a certain amount of the material is leached away, and Goldschmidt estimated that the mass of the sedimentary deposits is 0.97 of the original igneous rocks that gave rise to them.

Let X be the amount of igneous rock eroded per square centimeter of earth's surface. Let Y be the amount of clastic sediments deposited per square centimeter of earth's surface. Then, $Y = 0.97X$.

$$\text{Sodium content of igneous rock per cm}^2 = \frac{2.83}{100} \times X$$

$$\text{Sodium content of clastic rock per cm}^2 = \frac{1}{100} \times Y$$

However, the sodium content of ocean water per square centimeter = 2.975 kg. Therefore,

$$\frac{2.83}{100} X - \frac{Y}{100} = 2.975$$

and $X = 160 \text{ kg/cm}^2$
$Y = 155 \text{ kg/cm}^2$

The value of 160 kg/cm² gives a figure of about 3×10^8 km³ for the total amount of igneous rock weathered during geological time, a value close to that obtained by Clarke.

These calculations overlook two factors: the sodium removed from the ocean during geological time in the form of salt deposits and that added by way of volcanic exhalations and in solution in magmatic waters. Both these items are probably small in relation to the sodium cycle as a whole, and they work in opposite directions. In addition, of course, the calculations fail to take into account the sodium that may have been present in the primitive ocean; any such sodium would reduce the amount of weathering necessary to produce the present sodium content of sea water.

Goldschmidt also calculated the quantity of calcium and magnesium carbonate in sedimentary rocks. He estimated the average content of noncarbonate CaO in sandstones and shales to be 0.6% and concluded therefore that CaO in excess of this figure is present as calcium carbonate. Similarly, the average amount of noncarbonate MgO was estimated to be 2.6%. A balance sheet for the cycle of calcium and magnesium in sedimentation was then derived as follows:

In the ocean Ca $0.00042 \times 278 = 0.117 \text{ kg/cm}^2$
Mg $0.00130 \times 278 = 0.361 \text{ kg/cm}^2$

	CaO	MgO
In igneous rocks	5.08%	3.49%
In carbonate-free sediments	0.6%	2.6%
160 kg igneous rocks contain	8.128 kg	5.584 kg
155 kg carbonate-free sediments contain	0.930 kg	4.030 kg

278 kg sea water contain   0.117 kg Ca   —0.361 kg Mg
    corresponding to   0.164 kg CaO   0.598 kg MgO

Therefore, the following amounts of calcium and magnesium carbonates (per square centimeter) must be present in the sediments:

$$
\begin{array}{ll}
7.034 \text{ kg CaO} & 0.956 \text{ kg MgO} \\
\underline{5.519 \text{ kg CO}_2} & \underline{1.043 \text{ kg CO}_2} \\
12.553 \text{ kg CaCO}_3 & 1.999 \text{ kg MgCO}_3
\end{array}
$$

Assuming that all the magnesium carbonate in sediments is there as dolomite, $CaMg(CO_3)_2$, we obtain these figures: 10.170 kg $CaCO_3$, 4.372 kg $CaMg(CO_3)_2$. Hence the total amount of sedimentary rocks per square centimeter of the earth's surface is 155 kg clastics, 10.2 kg limestone, and 4.4 kg dolomite; the combined CO_2 in sediments is 6.562 kg/cm². From the above figures and the densities of these rock types the total average thickness of sediments can be calculated.

	Mass/cm²	Density	Volume (cm³)	Thickness
Clastics (sandstones and shales)	155 kg	2.65	58,491	585 m
Limestone	10.2 kg	2.7	3,777	38 m
Dolomite	4.4 kg	2.9	1,517	15 m
	169.6 kg			639 m

Clarke's calculation for the total sedimentary rocks of the globe gave an approximate thickness of 2500 feet or 762 meters. The agreement between the figures of Goldschmidt and Clarke is good, considering the assumptions involved in both cases.

Goldschmidt's computations have been extended and refined by later investigators. The resulting estimates for the relative amounts of the common sediments are collected in Table 6.7. The later results do not differ significantly from those of Goldschmidt, but the calculated total mass has increased considerably. The importance

Table 6.7   Estimates of the Relative Amounts of the Common Sediments, Based on Geochemical Calculations

	1	2	3	4	5
Shale	91	83	74	81	74
Sandstone		8	17	11	11
Limestone and dolomite	9	9	9	8	15
Total mass ($\times 10^{21}$g)	865	1066	2050	2500	3200

1. Goldschmidt, 1933
2. Wickman, 1954
3. Horn and Adams, 1966
4. Garrels and Mackenzie, 1971
5. As 4, but including allowance for submarine alteration of volcanic rocks.

of calculations of this kind lies not so much in precisely determining the amounts of each sedimentary class but rather in elucidating the general features of the geochemical cycle.

The results of the geochemical computations are in some disagreement with field measurements of the relative percentages of sedimentary rock types. Measurements of stratigraphic sections consistently give greater percentages of limestone and dolomite than indicated from geochemical calculations; for example, Leith and Mead assembled the data for 708,000 feet of sedimentary rocks in North America, Europe, and Asia and obtained the following percentages: shale 46, sandstone 32, limestone and dolomite 22. Such compilations must be used with caution, since the identification of a formation as a limestone, a sandstone, or a shale is liable to be very gross; shales often contain considerable sand, sandstones may carry much clay, and the term limestone is applied to many rocks with only 50% or so of carbonate. Nevertheless, it is clear that carbonate rocks are more prominent in the geological column than one might expect from geochemical calculations based on their derivation from continental igneous rocks. Garrels and Mackenzie have discussed this problem and present evidence for the derivation of much of the material of the carbonate sediments from the alteration of submarine volcanic rocks. The results of their calculations (Table 6.7, column 5) clearly provide better agreement with the observed proportions of sedimentary rocks.

Garrels and Mackenzie also discuss the variation in the proportions of the different types of sedimentary rocks over geological time. They note that the ratios of sedimentary rock types show striking changes over the past 2×10^9 years. Carbonate rocks make up 25% of the sedimentary rock mass of Phanerozoic (Cambrian-Recent) age, but only about 5% of rocks 600 million years to 2000 million years old. Evaporites are roughly 5% of the Phanerozoic rocks, but much less then 1% of older rocks.

SELECTED REFERENCES

Bathurst, R. G. C. (1975) *Carbonate sediments and their diagenesis* (second ed.). 658 pp. Elsevier Publishing Co., Amsterdam.

Bear, F. E. (ed.) (1964). *Chemistry of the soil* (second ed.). 515 pp. Reinhold Publishing Corp., New York. A comprehensive monograph by several authors on the different aspects of soil geochemistry; the article on trace elements by R. L. Mitchell is very informative.

Berner, R. A. (1971). *Principles of chemical sedimentology.* 240 pp. McGraw-Hill Book Co., New York. The application of thermodynamics and kinetics to the elucidation of problems of sedimentation, especially diagenesis.

Braitsch, O. (1971). *Salt deposits, their origin and composition.* 297 pp. Springer-Verlag, New York. A comprehensive discussion of salt deposits and the physical chemistry of their formation.

Degens, E. T. (1965). *Geochemistry of sediments.* 342 pp. Prentice-Hall Inc., Englewood Cliffs, N.J. A survey of the applications of low-temperature geochemistry to the study of sediments and sedimentary processes.

Feth, J. H., C. H. Roberson, and W. L. Polzer (1964). Sources of mineral constituents in water from granitic rocks, Sierra Nevada. *U.S. Geol. Survey Water-Supply Paper*, 1535-I. Weathering process of granite.

Garrels, R. M. and C. L. Christ (1965). *Solutions, minerals, and equilibria.* 450 pp. Harper and Row, New York. A comprehensive account of the principles of solution chemistry and thermodynamics as applied to mineral equilibria in sediments and sedimentary rocks.

Garrels, R. M. and F. T. Mackenzie (1971). *Evolution of sedimentary rocks.* 397 pp. W. W. Norton & Co., New York. A comprehensive account of the origin of sedimentary rocks and their development in space and time.

Goldich, S. S. (1938). A study in rock weathering. *J. Geol.* **46**, 17–58. A carefully documented account of the chemical and mineralogical changes produced by weathering on a granite gneiss, two diabases, and an amphibolite.

Grim, R. E. (1968). *Clay mineralogy* (second ed.). 596 pp. McGraw-Hill Book Co., New York. An authoritative work on the structure, composition, properties, occurrence, and mode of origin of the various clay minerals.

Harris, R. C. and J. A. S. Adams (1966). Geochemical and mineralogical studies on the weathering of granitic rocks. *Amer. J. Sci.* **264**, 146–173.

Horn, M. K. and J. A. S. Adams (1966). Computer-derived geochemical balances and element abundances. *Geochim. Cosmochim. Acta* **30**, 279–298. A new approach to the determination of elemental abundances in the crust and in major sedimentary rock types, using a computer program.

Imbrie, J. and A. Poldervaart (1959). Mineral compositions calculated from chemical analyses of sedimentary rocks. *J. Sed. Petrol.* **19**, 588–595. This paper provides a carefully reasoned procedure for obtaining a quantitative mineralogical composition from the chemical analysis of a sedimentary rock.

James, H. L. (1966). Chemistry of the iron-rich sedimentary rocks. *U.S. Geol. Surv. Prof. Paper 440-W.* 61 pp. [Data of Geochemistry (sixth ed.). chapter W.] A comprehensive account of the composition, distribution, and geochemistry of ironstones and iron-formation.

Krauskopf, K. B. (1955). Sedimentary deposits of rare metals. *Econ. Geol.*, 50th Ann. Vol., 413–463. An exhaustive compilation of the data on the occurrence and amounts of rare metals in different kinds of sediments and sedimentary rocks and a discussion of the physicochemical factors involved.

———(1956). Dissolution and precipitation of silica at low temperatures. *Geochim. Cosmochim. Acta* **10**, 1–26. An account of experimental work on the solubility of silica at low temperatures and the geological applications of this work.

Krumbein, W. C. and R. M. Garrels (1952). Origin and classification of chemical sediments in terms of pH and oxidation-reduction potentials. *J. Geol.* **60**, 1–33. A careful study of the pH and oxidation potentials of the common sedimentary environments and the resultant mineralogy.

Loughman, F. C. (1969). *Chemical weathering of the silicate minerals.* 154 pp. American Elsevier, New York. A short but comprehensive review of chemical weathering.

Pettijohn, F. J. (1975). *Sedimentary rocks* (third ed.). 628 pp. Harper & Row, New York. An excellent source of information on all aspects of sedimentary rocks,

especially useful because of the carefully chosen groups of analyses of different rock types.

Ronov, A. B. and A. A. Yaroshevsky (1969). Chemical composition of the earth's crust. In P. J. Hart (ed.), *The earth's crust and upper mantle, 37–57. American Geophysical Memoirs* **13**. Contains a detailed presentation of volume, mass, and average chemical composition of continental sedimentary and volcanic rocks and the oceanic sediments.

Ronov, A. B. and A. A. Yaroshevsky (1976). A new model for the chemical structures of the earth's crust. *Geochemistry International* 13, no. 6, 89–121.

Stewart, F. H. (1963). Marine evaporites. 53pp. *U. S. Geol. Surv. Prof. Paper 440-Y. [Data of Geochemistry* (Sixth ed.), chapter Y.] A brief but comprehensive account, both physiochemical and geological, of marine salt deposits.

Weaver, C. E. and L. D. Pollard (1973). *The chemistry of clay minerals.* 213 pp. American Elsevier, New York. A comprehensive account of the clay minerals in sediments and sedimentary rocks.

Wickman, F. E. (1954). The "total" amount of sediments and the composition of the "average igneous rock." *Geochim. Cosmochim. Acta* 5, 97–110. A careful evaluation of the data available for the calculation of the amount of igneous rock weathered and the amount of sediments formed during geological time.

7 ISOTOPE GEOCHEMISTRY

The study of the isotopes of some chemical elements has become one of the most important and powerful tools in geochemistry. Variations in the relative abundances of the isotopes of an element may result from several types of geochemical reactions. The measurement and understanding of these changes in many cases give clues to these reactions. Elements usually selected for isotopic investigations include those with two or more isotopes that have measurable isotopic abundance differences. Mass spectrometers used for the sensitive analyses are capable of measuring variations in the isotopic abundances with precisions in the vicinity of 0.01%.

Isotopic variations may be divided into two major groups: (a) those due to the radioactive decay of unstable nuclides, which causes variations in the isotopic composition of the stable daughter products and (b) those due to variations in nonradiogenic isotopes most often produced by exchange reactions, kinetic reactions in biological systems, or physical-chemical processes such as evaporation or diffusion. Radiogenic isotopes most often studied are those of argon, strontium, lead, and neodymium. These elements may show significant variations due to the length of time their parent radioactive nuclides have had to decay and the original concentrations of the parent and daughter nuclides. Strontium isotopes have proven to be a particularly useful tracer to indicate whether the magma that formed an igneous rock originated in the crust or the mantle. Magma derived from the upper mantle shows a lower concentration of ^{87}Sr due to the lower concentration of the radioactive parent ^{87}Rb in the mantle than in the crust. Stable isotope variations are most pronounced in the low atomic weight elements hydrogen, carbon, oxygen, and sulfur. Note that these elements may all be found in volatile or fugitive species as well as in solid phases. Stable isotope variations occur and are most easily detected when the relative

187

mass differences between the isotopes are relatively large, such as ^1H and ^2H or ^{12}C and ^{13}C with a mass differences of 100% and 8%. Isotopic variations are more likely to occur in the low atomic weight elements than those of higher atomic weight. With increasing temperatures the fractionation of the stable isotopes decreases, thus they show greater fractionation in sedimentary systems than in magmatic rocks. In general the degree of stable isotope fractionation decreases from the materials formed at the surface of the earth to those found at depth. Recycling of surficial material through the upper mantle tends to rehomogenize the isotopes.

STABLE ISOTOPES

As noted above, non-radiogenic variations in the relative isotopic abundances of several elements have been detected in nature. For the elements H, C, O, and S the observed differences are of geologic interest. The variations measured by mass spectrometry may be reported as absolute ratios such as $^{18}O/^{16}O$, but they are more often reported in terms of a factor δ in per mil (‰) or parts per thousand, where:

$$\delta = \left(\frac{R \text{ sample}}{R \text{ standard}} - 1 \right) 1000$$

and R is the ratio of two selected isotopes in a sample or standard. The use of this type of factor minimizes systematic errors in measurement between different spectrometers and laboratories. In Table 7.1 the measured isotope pairs for the commonly measured elements are given, together with the naturally observed isotopic ranges. For oxygen the per mil variation is for $^{18}O/^{16}O$ and the standards commonly used may be either standard mean ocean water, SMOW, as defined by Craig (1961), or PDB, the carbonate from a Cretaceous belemnite first used as a standard by Urey

Table 7.1

Element with Isotopic Standard Used	Isotopes Measured	Per Mil Isotopic Variation, δ of Heavier Isotope	
		Minerals and Rocks	Natural Waters
Hydrogen			
Standard mean ocean water (SMOW)	$^2H/^1H$ (D/H)	−180 to +20	−410 to +50
Carbon			
Calcite-PDB (belemnite)	$^{13}C/^{12}C$	−35 to +5	
Oxygen			
SMOW	$^{18}O/^{16}O$	−2 to +36	−50 to +15
Silicon			
Quartz vein, Mother Lode, California	$^{30}Si/^{28}Si$	−2.2 to +3.2	
Sulfur			
Canyon Diablo meteorite troilite	$^{34}S/^{32}S$	−45 to +60	

et al. (1951). Likewise $^2H/^1H$ variations are referred to SMOW and $^{13}C/^{12}C$ variations to PDB. Troilite (FeS) from the Canyon Diablo meteorite is the standard used for $^{34}S/^{32}S$. Several of the original standards have been essentially exhausted, but laboratories today may use secondary standards calibrated against these original materials. Positive or negative δ values indicate enrichment or depletion of the heavier isotope in the sample relative to the standard.

The isotopic variations are due to mass differences between the isotopes that lead to small differences in the vibration frequencies of the atoms in a molecule or crystal. As a result of the differing vibration frequencies, the thermal properties of the isotopes, such as internal energy (E), heat capacity, and entropy (S), vary and may lead to a fractionation in the distribution of the isotopes between two phases. For the distribution of ^{18}O and ^{16}O, for example, a separation constant α between two phases may be calculated as

$$\alpha = \frac{R_A}{R_B} = \frac{(^{18}O/^{16}O) \text{ phase } A}{(^{18}O/^{16}O) \text{ phase } B}$$

In equilibrium systems it may be shown that α is the equilibrium constant of a isotopic exchange reaction if the reaction is written for the exchange of one atom. Such an equilibrium constant may be calculated using statistical mechanics or measured for a number of phases of geochemical interest. The constant is temperature-dependent and hence a useful natural thermometer for geochemical systems. The fractionation factor α is related to δ by the relations:

$$\alpha = \frac{1 + \delta_A/1000}{1 + \delta_B/1000}, \qquad 1000 \ln \alpha \simeq \delta_A - \delta_B, \qquad \text{and } \alpha \simeq 1 + \frac{\delta_A - \delta_B}{1000}$$

For biological systems it is difficult to precisely predict fractionations produced by rate-controlled or diffusion reactions, but they may be measured in typical systems. Generalizations that may be made about isotope fractionation include: (a) bonds formed by light isotopes are more easily broken than those of heavy isotopes; (b) molecules with light isotopes will react more readily than those with heavy isotopes; and (c) lighter isotopes show preferential enrichments in irreversible chemical reactions. The measured isotopic variations are a function of the fractionation processes and the original isotopic composition of the starting materials. The results may give information on (a) the temperature of formation of rocks and minerals, including fossils, (b) chemical and physical processes that have affected rocks during or following their formation, and (c) genetic relations among rock and meteorite types.

OXYGEN ISOTOPES

Oxygen is the most abundant element in the earth's crust and hence many oxygen-containing systems are available for study. The mass spectroscopic measurements are most often made on oxygen in the form of CO_2. Figure 7.1 shows $\delta^{18}O$ decreases with increasing temperature in natural waters, and lower-temperature materials show larger variations in $\delta^{18}O$ than do the high-temperature igneous rocks.
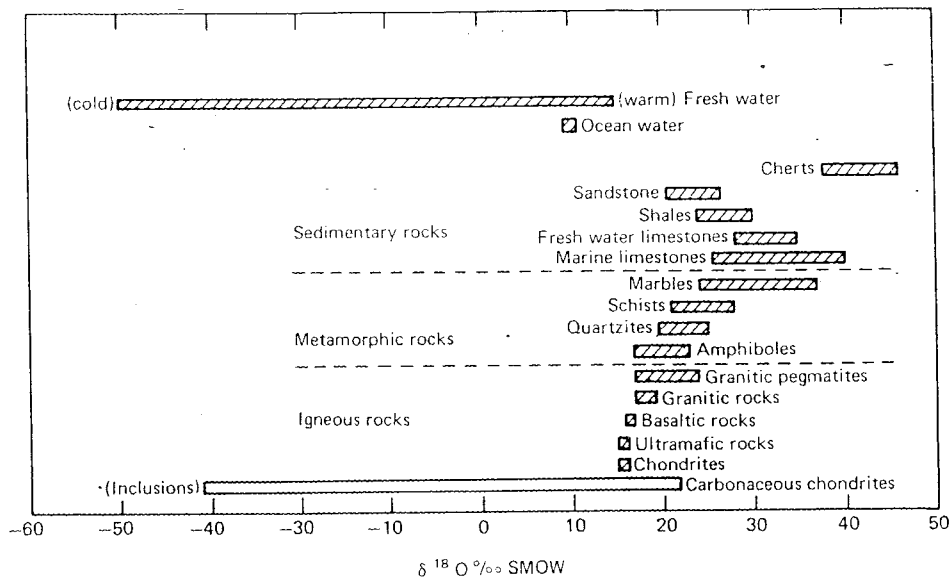
Figure 7.1 Oxygen isotope variation in common rocks, minerals, and waters.

The oxygen isotope fractionation in igneous rocks approaches equilibrium, and hence α values may be calculated and temperatures of formation determined. The measured values may be compared with experimentally determined fractionation curves as shown in Figure 7.2. In this figure mineral vs. water curves have been plotted. From such data, $\delta^{18}O$ differences between coexisting mineral pairs can be used to indicate temperatures of formation or a graph similar to that shown in Figure 7.3 may be determined and the temperatures estimated directly.

It may be noted that minerals in which Si—O—Si bonds predominate become enriched in ^{18}O. Bonds of the Si—O—Al type and Si—O—Mg types show less fractionation. Quartz shows the highest ^{18}O enrichment of the common rock-forming minerals, followed by the feldspars, muscovite, amphiboles and pyroxenes, biotite, and olivine. The oxides magnetite and ilmenite show the least fractionation. The results obtained in many igneous rocks indicate that their minerals were in equilibrium with a large reservoir of constant $^{18}O/^{16}O$ ratio. The diffusion of oxygen in the crust and upper mantle appears to be relatively rapid.

SULFUR ISOTOPES

Meteoritic troilite has a remarkably constant $^{34}S/^{32}S$ ratio of 22.21, and may perhaps be considered to represent primordial sulfur. The ^{34}S content of sulfur in components of the earth's crust varies by about 50 ‰ on either side of this standard, as illustrated in Figure 7.4. The variations, which appear large in view of the relatively high atomic weight of sulfur, are due in large part to the action of sulfate-reducing bacteria that produce sulfides enriched in ^{32}S whereby the sulfate remaining becomes enriched in

^{34}S. This mechanism can account for the increase of ^{34}S in oceanic sulfate residual after bacterial reduction. The H_2S produced is removed from the ocean by precipitation as iron and other heavy-metal sulfides. Processes such as the oxidation and reduction reactions of H_2S, SO_2, and S during volcanism, and sulfide-sulfate equilibrium under magmatic or hydrothermal conditions, may be locally important. Most hydrothermal sulfide deposits do not show great isotope variations, indicating the sulfur comes from a primitive reservoir or a homogenized system. Exotic values from selected deposits may indicate that hydrothermal sulfur is from a fractionated sedimentary source. Such sulfur may be light if derived from sedimentary sulfides or heavy if derived from sedimentary sulfate.

One of the more successful applications of sulfur isotopes has been in demonstrating that the elemental sulfur in the salt domes of Louisiana and Texas was formed
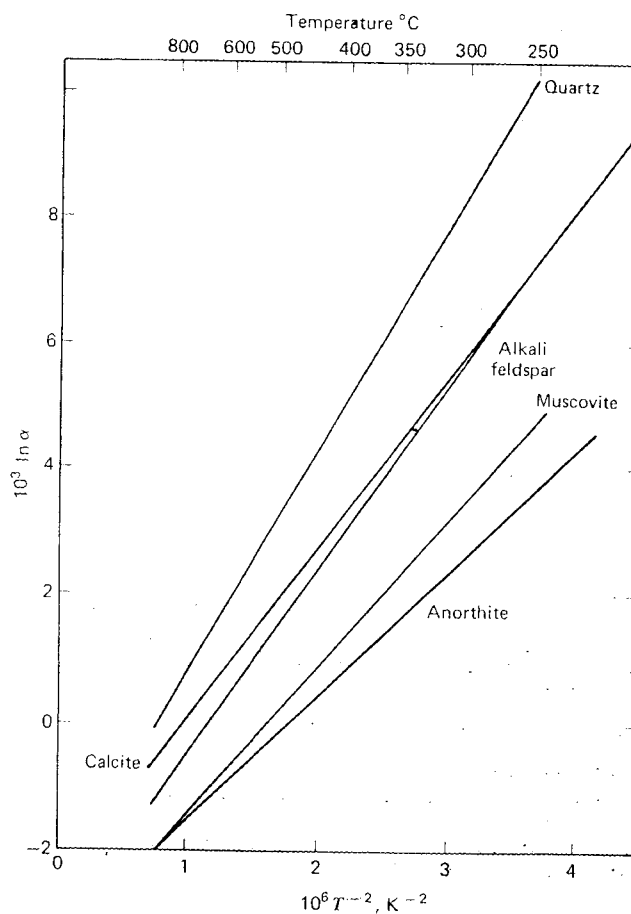
Figure 7.2 Experimentally determined oxygen isotope fractionation curves for selected mineral-H_2O systems.

by the bacterial reduction of sulfate in the gypsum or anhydrite cap rock. These salt domes are capped by anhydrite apparently precipitated in an inland sea. The $^{32}S/^{34}S$ ratio in this anhydrite is a relatively constant 21.85. The anhydrite was reduced by the bacteria *Desulfovibrio*, which used accumulating petroleum as their energy source. The petroleum was oxidized to CO_2 while the sulfate was being reduced to H_2S. The H_2S then reacted with residual sulfate to produce native sulfur. This mechanism is supported by isotopic measurements of H_2S in the salt dome cap rock, which has a 2 to 5% higher $^{32}S/^{34}S$, complementing the enrichment in ^{32}S cap rock hydrogen sulfide. Native sulfur in cap rock is slightly heavier than the associated hydrogen sulfide and lighter than the associated sulfate, verifying its production by the $SO_4 - H_2S$ reaction. The sulfur data are supported by the $^{13}C/^{12}C$ ratios of calcite cap rock, which are lower than those of sedimentary carbonate but are similar to those of
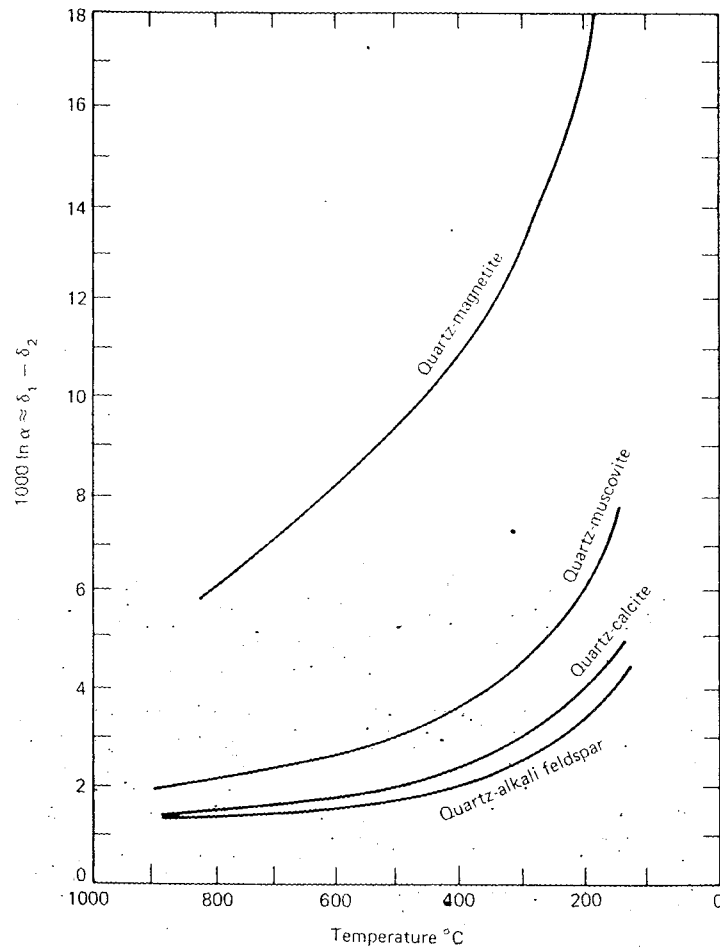
Figure 7.3 Oxygen isotope calibration curves for quartz-mineral systems.
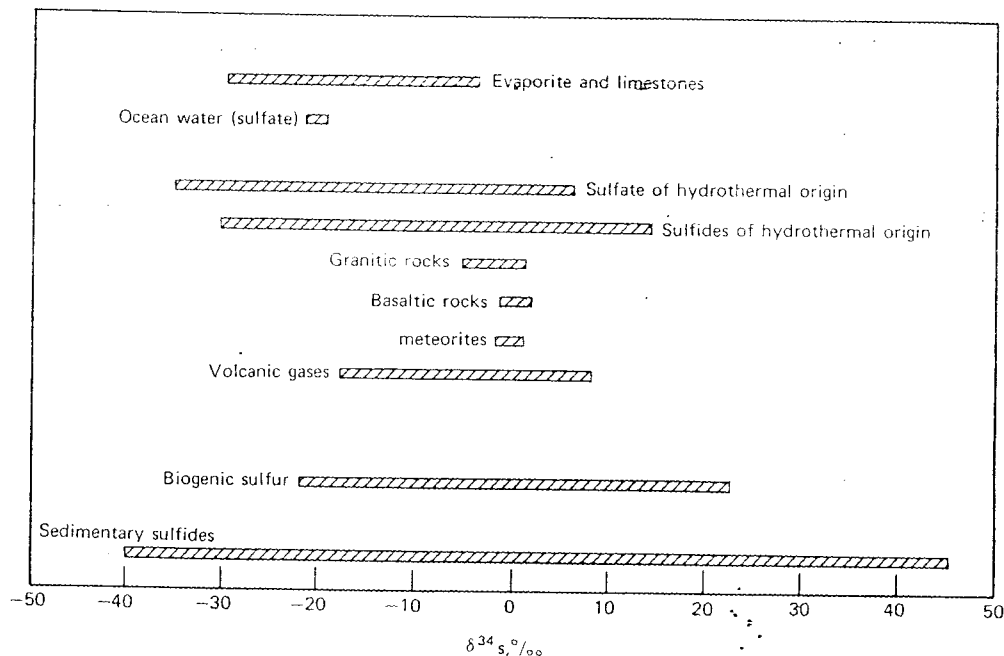
Figure 7.4 Sulfur isotope variation in common rocks, waters, and volcanic gases.

petroleum, thus indicating that the calcite carbon dioxide was derived from the oxidation of petroleum.

CARBON ISOTOPES

Carbon, like sulfur, shows large isotopic differences in geochemical samples. The variations may be ascribed to the action of organisms with isotopic fractionation caused by kinetic effects and also to the ability of carbon to exist in a large number of compounds in natural systems. The data in Figure 7.5 show that the heavier carbon isotope ^{13}C tends to be concentrated in carbon-oxygen inorganic compounds. Since the reference standard for carbon isotope measurements is a fossil carbonate shell, many of the reported measurements have negative $\delta^{13}C$ values. Several aspects of organic carbon isotope studies and additional carbon isotope data are included in Chapter 10. In addition to the stable isotopes of carbon ^{12}C and ^{13}C, an exceedingly small amount of the radioactive isotope ^{14}C may also be found. ^{14}C in natural systems is not detectable by mass spectrometry.

Meteorites exhibit large carbon isotope variations. This appears to be due to the presence of several carbon phases in the different types of meteorites. The carbonate phases in carbonaceous chondrites have $\delta^{13}C$ values near $+60‰$ whereas the coexisting hydrocarbon phases are around $-10‰$. The whole-rock value of $\delta^{13}C$ of $-7.2‰$ for the Murchison C2 carbonaceous chondrite may represent a reasonable value for total terrestrial carbon. Figure 7.6 shows that atmospheric CO_2, carbonatite, and diamond deposits also have carbon isotope values near this value.
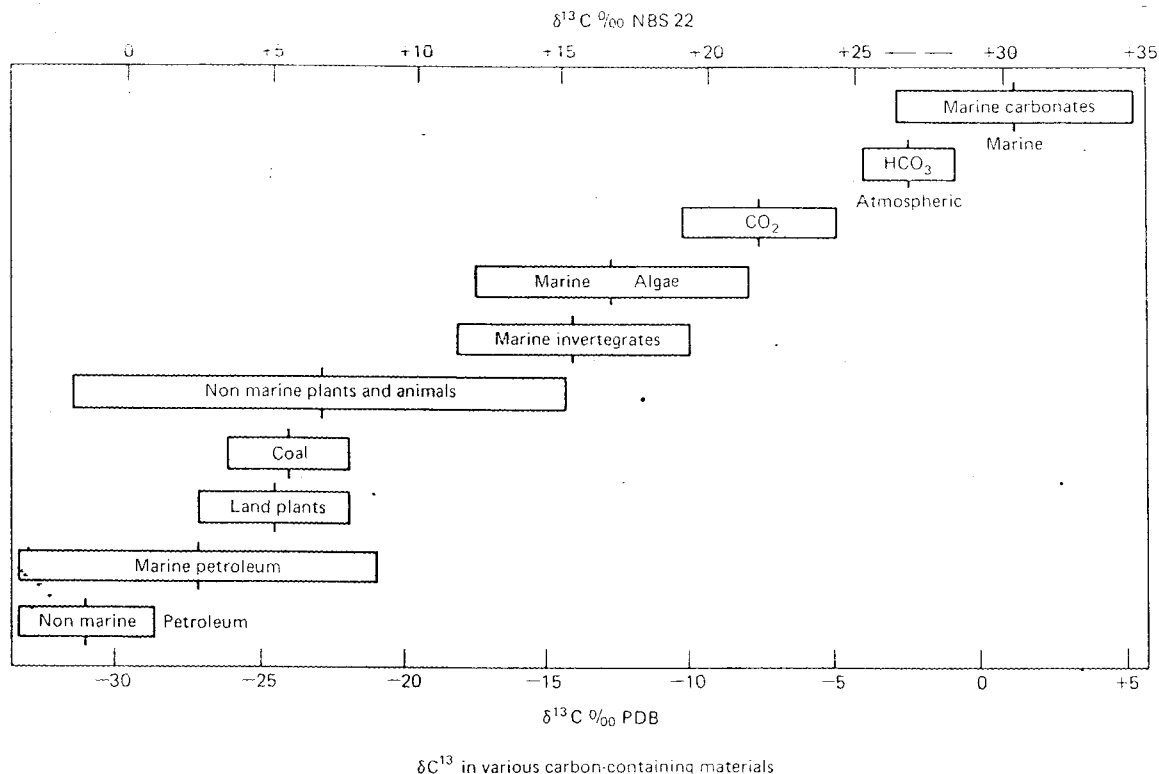
Figure 7.5   Carbon isotope variation in natural carbon-containing materials.

When chemical fractionation of carbon takes place in the geochemical cycle, isotopic fractionation also occurs. In volcanic and fumarole gases methane carbon is lighter than carbon dioxide carbon. Methane in Yellowstone and New Zealand fumaroles has an average $\delta^{13}C$ of $-28‰$ while associated CO_2 has an average $\delta^{13}C$ of about $-4‰$. Comparison of these values with those calculated for the reaction $CO_2 + 4H_2 = CH_4 + 2H_2O$ indicates that these gases equilibrated at about 200 to 300°C. These calculations compare well with bore-hole measurements, and, because of the slowness of the equilibration in the above system, indicate that the components existed at these temperatures for a long period of time.

Most inorganic carbon isotope fractionation is due to equilibration effects. The equilibrium of carbon dioxide in the atmosphere with bicarbonate in the oceans is described by the equation:

$$H^{12}CO_3^- \text{ (aq)} + {}^{13}CO_2 \text{ (gas)} = H^{13}CO_3^- \text{ (aq)} + {}^{12}CO_2 \text{ (gas)}$$
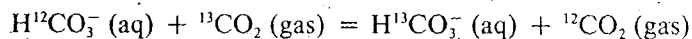
This reaction has an equilibrium constant of about 1.005 at 20°C, and the measured values $\delta^{13}C$ of -7 in atmospheric CO_2 and -2 in oceanic bicarbonate are in agreement with this. Likewise the equation $H^{13}CO_3^- \text{ (aq)} + Ca^{12}CO_3 = H^{12}CO_3 + Ca^{13}CO_3$, with an equilibrium constant of 1.004 at 20°C, predicts that if isotopic

equilibrium were attained, the marine carbonates would have a $\delta^{13}C$ about 4% heavier than the marine bicarbonate from which they were derived. Variations from this value in measured systems may indicate that biochemical metabolic reactions cause kinetic isotope effects complicating this system. Any local variations in the $\delta^{13}C$ of natural waters will directly affect the $\delta^{13}C$ in carbonates precipitated from them. Such variations are more likely to be found in freshwater systems than in ocean water. Changes in carbon isotopic compositions over the geologic record may have occurred in natural reservoirs in the geologic past. Any real differences are most easily explained by variations in photosynthetic activity. Slight maximums in $\delta^{13}C$ ratios in Devonian and Cretaceous times are perhaps due to the development of land plants and flowering plants in these times. The major difficulties in interpreting such
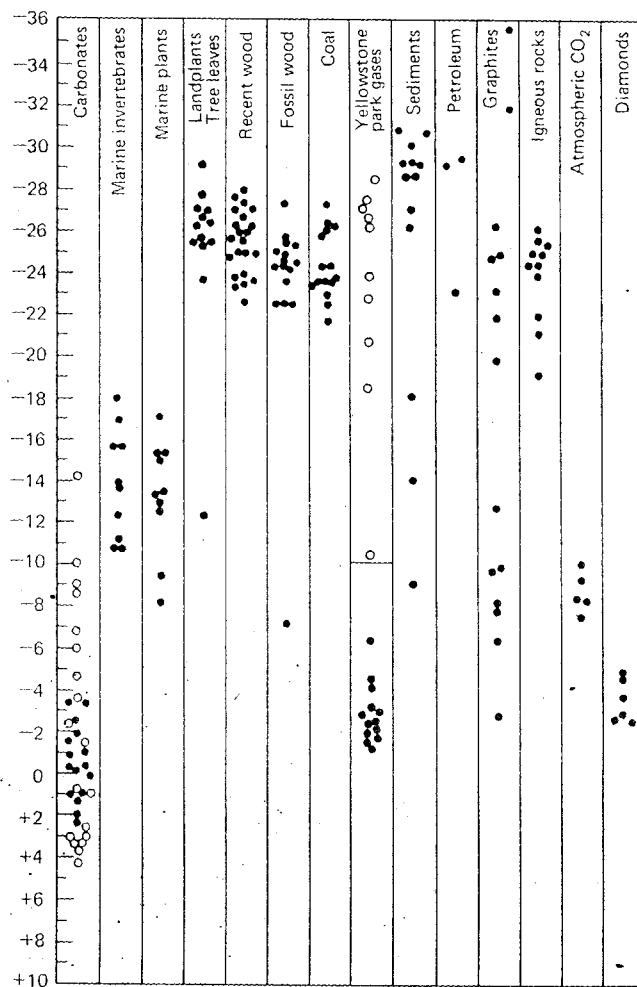
Figure 7.6   Isotopic abundances of carbon (Craig, 1953, *Geochim. Cosmochim. Acta*, 3, p.60).

changes come from our inability to evaluate slow secondary exchange reactions in these samples over the long time periods since their deposition.

HYDROGEN ISOTOPES

Because the relative mass difference between the two stable isotopes of hydrogen ^1H and ^2H (deuterium, D) is larger than any other isotope pair, hydrogen shows the greatest isotopic fractionation. Fractionation up to 70% or 700‰ has been recorded on geochemical samples. The δD values for common geochemical materials are shown in Figure 7.7. The standard for the deuterium measurements is the same SMOW used for oxygen isotopes. Note in Figure 7.7 that most samples have δD values lighter than the ocean water, which is essentially constant in its hydrogen isotope composition. Since the vapor pressure of HDO is lower than H_2O, evaporation mechanisms thus remove the isotopically lighter molecule from ocean water. This fractionation is paralleled by a similar and smaller fractionation of oxygen in the same waters. Measurements of precipitation show that waters near the equator have both hydrogen and oxygen isotopic compositions near that of ocean water, whereas snow near the poles may have D values of less than $-300‰$ and $\delta^{18}O$ values in the range of $-45‰$ to $-50‰$. This fractionation is quite regular with respect to latitude and may be explained by assuming the earth's atmosphere is acting as a giant condenser. As water evaporates or condenses D and ^{18}O are concentrated in the liquid phase. The relationships between δ^2H, (δD) and $\delta^{18}O$ are shown in Figure 7.8. Only rarely does the isotope composition of natural waters fall away from this line. The variations therefore provide unique fingerprints for meteoric waters. Meteoric waters that become groundwaters exchange their oxygen isotopes with carbonate and silicate minerals that they come in contact with. The hydrogen isotopes, however, show minimal exchange since there is relatively little hydrogen in the rocks the water may pass through. One of the principal conclusions drawn from stable isotope studies of waters in geothermal systems is that most such waters are not primary magmatic waters but meteoric waters from local precipitation. This is illustrated in Figure 7.9
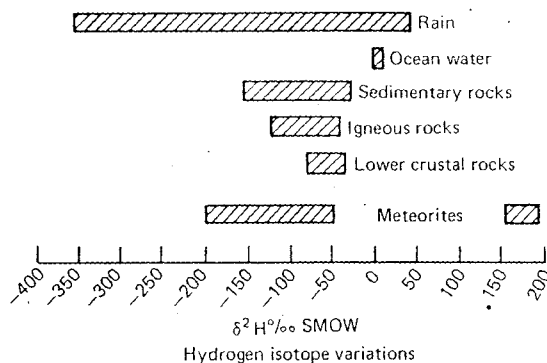
Figure 7.7   Hydrogen isotope variations.
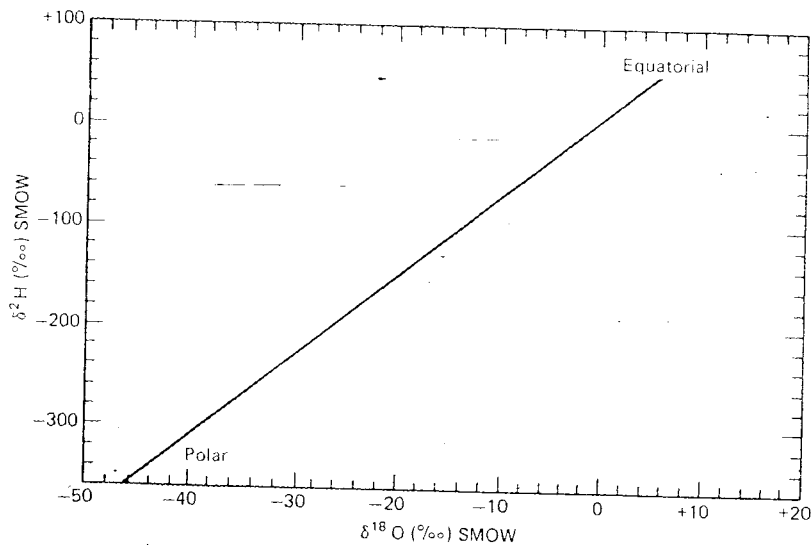
Figure 7.8   Isotopic variations in meteoric water.
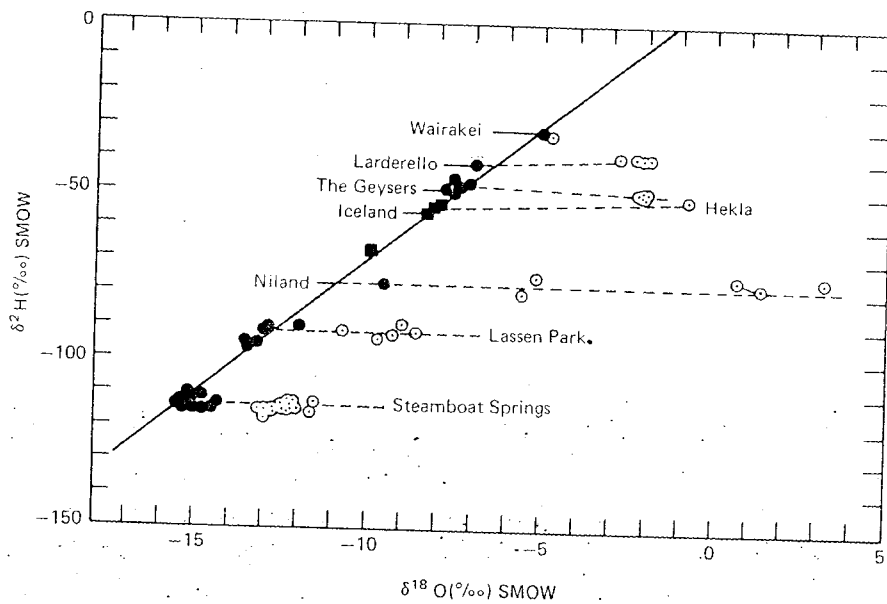
Figure 7.9   Isotopic variations in some geothermal waters and in geothermal steam. ■: local meteoric waters or slightly heated nearsurface ground waters; ● : hot springs or geothermal water; ⊙ : high temperature, high pressure, geothermal steam. Niland = Salton Sea Geothermal Area. (After CRAIG, 1963. *Conf. on Nuclear Geology in Geothermal Areas.* Spoleto, Italy.)

in which the δD and $\delta^{18}O$ values for local meteoric waters and related geothermal waters are plotted. Although the $\delta^{18}O$ values in a given area may change by isotopic exchange with local rocks, the δD values remain constant, supporting a meteoric origin for the hydrogen. If it were diluted by primordial hydrogen the lines would be skewed rather than horizontal. The same effect is also noted where nonequlibrium evaporation takes place, as in the acid type hot springs of Yellowstone National Park.

OTHER ISOTOPES

Isotopic investigations of H, C, O and S have shown them to be the most useful elements for elucidating geochemical systems. Other elements have been studied with less success. Nitrogen isotopes in natural organic compounds show variations in $\delta^{15}N$ of about 10‰ above and below the standard atmospheric nitrogen. The mechanisms for producing these changes are not well understood. Likewise silicon isotopes show variations in $\delta^{30}Si$ of about 2‰ in igneous and sedimentary rocks. Other elements of minor interest, such as selenium and boron, have measurable isotopic variations. The changes in selenium are attributed to biochemical activity, and boron to selective incorporation of ^{10}B into clay minerals as compared to the heavier ^{11}B isotope.

GEOCHRONOLOGY

An interesting and widely used application of mass spectrometric isotope studies has been in the absolute dating of geologic events. In order to use isotope studies for geochronologic purposes, careful studies of radiogenic daughter products of radioactive nuclides and the concentrations of the parent radioactive nuclide must be made. Such systems are known as accumulation clocks. In addition it has also been possible to directly study the decay of short-lived radioactive nuclides, notably ^{14}C. These methods are known as decay clocks. In either method the precision of the geologic date determined depends on instrumental sensitivity, sample handling, and sample properties to be discussed later.

The radioactive nuclides commonly used in geochronology are given in Table 7.2. Note that all of these are accumulation clocks except for ^{14}C.

RADIOACTIVE DECAY

Radioactive decay may take place in several ways. In the systems commonly used for geochronology the uranium and thorium isotopes decay through a series of α and β decays until they change into stable lead isotopes. The details of these decay schemes are given in appendix II. Rubidium-87 changes by β decay into strontium-87 and carbon-14 changes by β decay to nitrogen-14. Potassium-40 may decay by one of two alternate routes: by β decay to calcium-40 or K electron capture to form argon-40. The argon-40 daughter product is the one used to date geologic samples.

Table 7.2

Nuclides	Half-Life (yr)	λ (yr^{-1})	Effective Range (yr)	Materials
$^{235}U-^{206}Pb$	4.47×10^9	1.55×10^{-10}	10^7-10^9	zircon, uraninite
$^{235}U-^{207}Pb$	0.71×10^9	9.72×10^{-10}	10^7-10^9	zircon, uraninite
$^{232}Th-^{208}Pb$	1.39×10^{10}	4.99×10^{-11}	10^7-10^9	zircon, monazite
$^{87}Rb-^{87}Sr$	4.88×10^{10}	1.42×10^{-11}	10^7-10^9	micas, igneous and metamorphic rocks
$^{147}Sm-^{143}Nd$	1.06×10^{11}	$6/54 \times 10^{-12}$	10^9	igneous rocks
$^{40}K-^{40}Ar$	1.31×10^9 (total)	$\beta\ 4.72 \times 10^{-10}$ $\kappa\ 5.85 \times 10^{-1}$	10^4-10^9	micas, hornblend, sanidine, igneous and metamorphic rocks
^{14}C	5730	1.21×10^{-1}	$0-10^5$	wood, charcoal, bone, shells

The usefulness of radioactive decay for dating old events is based on the assumption that the rate of decay is constant with respect to time and chemical state, temperature, pressure or other environmental factors of the radioactive parent isotope. The rate of decay of a radioactive nuclide is a first-order kinetic reaction. If N is the number of atoms present at a given time and λ is the decay constant for the radioactive nuclide the loss of N with time is $-dN/dt = \lambda N$. Integration of this equation and letting N_0 be the initial number of radioactive atoms yields the relationship

$$\frac{N}{N_0} = e^{-\lambda t}$$

From this equation the relationship between the half-life of a radioactive nuclide and the decay constant may be derived by substitution ½ for N/N_0 which gives $t_{half-life} = 0.693/\lambda$.

Since most of the radioactive dating techniques used in geochronology are long half-lived accumulation clocks, it is convenient to put the decay equations in terms of N_D the atoms of daughter nuclide and N the atoms of parent nuclide present at the time an age date is measured. The relationship is

$$N_D = N(e^{\lambda t} - 1)$$

If the half-life is long with respect to the measured age the relationship may be simplified by the approximation

$$\frac{N_D}{N} = \lambda t$$

The fact that the decay of the uranium isotopes to lead isotopes takes place by a series of sequential decays is very useful in geochronology in that it allows the half-lives of the long-lived radioactive isotopes to be measured with more accuracy than they could by direct measurement. In these sequences where intermediate nuclear species are in radioactive equilibrium or being formed and decaying at the same rate, the decay constants and nuclide abundances are related by $\lambda_1 N_1 = \lambda_2 N_2 = \lambda_3 N_3 = \ldots$. Hence the direct measurement of the half-life or decay constant for a short-lived intermediate allows the decay constant of the long-lived rate-determining decay to be calculated.

For a radioactive isotope to be useful for geochronology, several criteria must be satisfied. Either the radioactive isotope must be continually replenished as in the case of ^{14}C, or the half-life of the radioactivity must be about the age of the earth so that significant amounts remain. The half-life must not be too long or the production of daughter product will be too low. The parent nuclide must be moderately abundant in rocks and significant detectable enrichment of the daughter nuclide must take place. In addition the rate of decay of the parent must be known and the initial content of the parent and/or the daughter must be determined. Also the system must be closed such that no parent or daughter isotopes are added or subtracted by any process other than radioactive decay, or such change must take place in a regular and orderly manner.

Uranium-Thorium-Lead Methods

The uranium-thorium-lead system was one of the earliest accumulation clocks used in geochronology. The uranium and thorium were concentrated in uranium-rich minerals, often in pegmatite deposits, and were sometimes accompanied by lead-bearing sulfide minerals, so that corrections for the isotopic composition of the original lead could be determined. The presence of three separate but related radioactive clocks in the same system allowed checks to be made with respect to the basic premises of geochronology. If the three dates obtained were similar the results were termed concordant. If different dates were found they were called discordant. While in the early days of radioactive geochronology it was always hoped that concordant dates would be obtained, we now have developed our understanding of the systems involved to sometimes successfully use discordant dates to elucidate the history of a rock mass. The requirement of the pegmatitic uranium minerals greatly limited the scope of the early work. Today most uranium-lead dating is done on zircon crystals in common igneous rocks. Seldom are samples other than igneous rocks dated by accumulation clocks.

Natural uranium isotopes ^{238}U and ^{235}U decay through a series of intermediates to form ^{206}Pb and ^{207}Pb. The single Th isotope ^{232}Th decays to ^{208}Pb. The intermediate products in the decay chain have low concentrations and are not the cause of real errors in the calculated radioactive dates. Diffusion of an intermediate nuclide may cause calculated ages to be anomalous.

For minerals older than several million years, the radioactive dates for the three systems may be calculated by the equations

$$t_{206} = \frac{1}{\lambda_{238}} \ln \left(1 + \frac{^{206}Pb}{^{238}U} \right)$$

$$t_{207} = \frac{1}{\lambda_{235}} \ln \left(1 + \frac{^{207}Pb}{^{235}U} \right)$$

$$t_{208} = \frac{1}{\lambda_{232}} \ln \left(1 + \frac{^{208}Pb}{^{232}Th} \right)$$

In addition a lead-lead age may be calculated from the equation

$$\frac{^{207}Pb}{^{206}Pb} = \frac{^{235}U}{^{238}U} \cdot \frac{(e^{\lambda_{235}t} - 1)}{(e^{\lambda_{238}t} - 1)}$$

Natural uranium has a $^{235}U/^{238}U$ ratio of $1/137.8$ and this can be used in the above equations. In the lead-lead equation it is not necessary to measure the absolute quantity of lead or uranium but just the radiogenic $^{207}Pb/^{206}Pb$ ratio.

In order to correct for the original concentrations of lead isotopes use is made of the nonradiogenic ^{204}Pb. If ^{204}Pb is present in a uranium mineral, the presence of original lead is confirmed and must be corrected for. The best way to do this is by direct measurement from an adjacent lead-bearing, uranium-thorium-free mineral formed at the same time as the mineral to be dated. As a second choice the isotopic composition of lead from the general area may be used for the correction.

In an ideal closed system a plot may be made showing the locus of pairs of the ratios of $^{206}Pb/^{238}U$ and $^{207}Pb/^{235}U$ that have the same age. Such a plot as shown in Figure 7.10 is known as a concordia plot. It illustrates the expected ratios in closed-system samples of different ages as marked on the curve. Few actual data fit on such a curve but rather fall on lines similar to the example also shown in Figure 7.10. The distribution of the data points for discordant data is interpreted to be due to the loss of radiogenic lead from the minerals. In many such discordant data plots the data points fall on straight lines, which may be extrapolated back in time to intersect the concordia curve. The point of intersection with the concordia curve is the starting time of the radioactive clock. At one time the intersection of the straight line at the lower end of the concordia curve was interpreted to be the episodic time of lead loss. Further study has indicated that although such a process would give the observed distribution of data points, a similar distribution may be obtained by a constant or increasing diffusion of lead throughout the age of the rock in many cases. Hence it is usually the upper intersection that proves to be of greatest interest and use.

An indication of the age of lead deposits may be obtained by measuring the lead isotope contents. By assuming that the composition of lead in iron meteorites is the same as the composition of primordial lead at the time the earth's crust evolved, a lead evolution diagram may be constructed showing the evolution of lead isotopes over geologic time. In order to construct such a diagram an assumption must be made with respect to the ratio of uranium and lead present in the earth. Figure 7.11 shows a lead evolution diagram for a system with a present-day $^{238}U/^{204}Pb$ ratio of 9.

Measurements of lead from some ore deposits give concordant ages and fall on such a smooth curve. This indicates that they have been extracted at a given time from a region of average U-Pb distribution and hence have the composition of ordinary or common lead at that point in the earth's history. There are however, many lead samples that isotopically diverge from this regular system. These so-called anomalous Pb systems are explained by the contribution of lead from local U-Th-Pb systems that have not been well mixed into a crustal average. An anomalous lead curve is also shown in Figure 7.11. The extension of the line above and beyond the common lead curve indicates that a uranium-rich source, such as granitic crustal rocks, has added excess radiogenic lead to the samples. It may be that thorough isotopic mixing may require transport of material into the upper mantle by subduction. Partial melting of crustal materials may produce anomalous lead. The study of lead isotopes may give much geologic insight into the origin of lead ores but great caution must be taken lest serious errors arise from false assumptions.
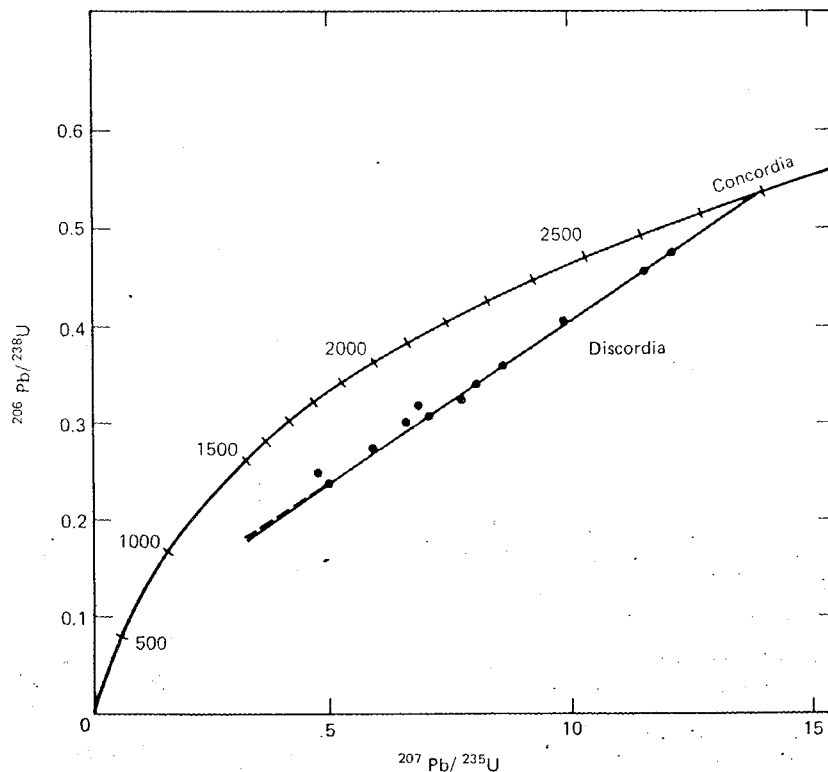
Figure 7.10 Concordia plot for the U-Pb system. Ages in years before present are marked on the Concordia curve. A typical set of measurements is shown by Discordia. The intersection of this line with Concordia is interpreted as the age of the sample.
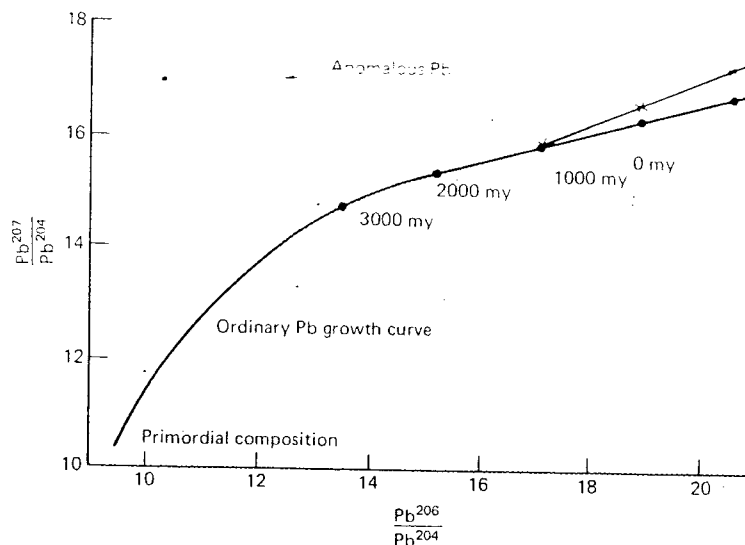
Figure 7.11   Diagram of Pb^{207}/Pb^{204} vs. Pb^{206}/Pb^{204} indicating the primordial isotopic compòsition and a growth curve for ordinary lead corresponding ìo a source region in which U^{238}/Pb^{204} (today) = 9-9. The isotopic composition of lead varies according to the time of isolation from the source region.

Potassium-Argon System

The decay of ^{40}K to ^{40}Ar has been one of the most widely used reactions for geologic age determination. It is especially useful in the age determination of rocks a few hundred thousand years old that are difficult to date by other techniques, as well as for older samples. This is primarily because argon is an inert gas and can be measured easily and accurately in small quantities. The major sources of error are old inherited argon and contamination by atmospheric argon during the analytical procedure. Potassium is one of the most abundant elements in the earth's crust and is found in many rock-forming minerals. The ^{40}K half-life of 1.31×10^9 years is optimal as compared to the age of the earth and the rate of decay. Argon does not enter into chemical combinations with other elements and only rarely is it retained directly in minerals as they form from magmas. Unfortunately radiogenic argon has a tendency to leak out of some minerals and care must be taken in selecting the proper rocks or minerals for age dating. Argon loss is increased by heating.

Potassium -40 may decay in two ways; by β decay to form ^{40}Ca and K-electron capture to form ^{40}Ar. Only the latter is extensively used because calcium is a common element in rocks and ^{40}Ca is the most abundant isotope of calcium. The branching decay requires a modification of the fundamental age relation equation to give

$$t = \frac{1}{\lambda} \ln \left[1 + \left(\frac{\lambda_\beta + \lambda_k}{\lambda_k} \right) \frac{^{40}Ar}{^{40}K} \right]$$

where λ is the decay constant of ^{40}K, λ_β is the rate of decay of ^{40}K to ^{40}Ca and λ_k is the rate of decay of ^{40}K to ^{40}Ar. The ratio of β decay to K-electron capture is about 8.1/1. If total potassium is determined in a dated sample it must be corrected to account for the fact that common potassium contains 0.0119 atom % ^{40}K. Correction for atmospheric contamination of argon can be corrected from the known atomic ratio of $^{36}Ar:^{38}Ar:^{40}Ar$ in air of 0.377:0.063:99.60. Most of the argon in the atmosphere has come from the decay of ^{40}K.

Studies have shown that the micas and amphiboles have the best argon retention of common rock-forming minerals, hence they are used most commonly for age determinations. Feldspars are rather poor argon retainers. With any of these minerals metamorphic alterations drive out the argon and a K-Ar determination on a metamorphic rock usually gives the date of the metamorphic event.

In K-Ar age determinations total K is usually determined by chemical methods. Argon is released from the sample by melting in a vacuum system and cleaned up by removing chemically active volatiles released with it. It is analyzed in a mass spectrometer and the abundance of radiogenic ^{40}Ar evaluated by adding a measured ^{38}Ar spike and also correcting for atmospheric contamination.

Care must be taken in the interpretation of K-Ar dates. They are best used in conjunction with the other dating methods and often indicate metamorphic events. For younger samples, the technique is invaluable and may be used for whole-rock analyses as well as mineral separates.

To avoid some of the problems related to Ar loss a technique known as the ^{40}Ar — ^{39}Ar method has been developed. It consists of converting a fraction of the ^{39}K in the rock to ^{39}Ar by neutron activation and then heating the sample in increments to release both the ^{39}Ar and ^{40}Ar. If there has been no loss of ^{40}Ar the proportion of released ^{39}Ar and ^{40}Ar will remain constant with increasing temperature and a date may be calculated. If a rock or mineral has lost radiogenic ^{40}Ar the released ratios will vary and indicate a date is unreliable. Under some circumstances the data may be saved if the interpretation can be made that the early release of argon comes from the open sites and the later-released argon is from sites that have good retention. Thus the high-temperature-released argon may yield a reliable age.

Rubidium-Strontium Method

The long half-lived decay of ^{87}Rb to ^{87}Sr has been particularly useful for dating relatively old geologic events. Although Rb is not a particularly abundant element, it is concentrated in K-rich minerals. The method has been used in minerals and rocks with high K contents. Corrections must be made for any primary ^{87}Sr incorporated in a rock or mineral upon solidification. This may be done by analyzing a nearby rubidium-free mineral for its strontium isotopic composition. The determination of this correction provides the major analytical uncertainty with the method.

Natural Rb contains 27.8 atom % of the radioactive ^{87}Rb and 72.2 atom % ^{85}Rb.

As expected the isotopic makeup of Sr may vary depending on its geochemical history. A long period in contact with ^{87}Rb will increase its proportion of ^{87}Sr. In analyses the isotopic composition and abundance of Sr is measured with a mass spectrometer. This method is most often used to determine the Rb abundance as well. After the radiogenic ^{87}Sr is determined the age of a rock may be found using the standard age relationship.

$$t = \frac{1}{\lambda} \ln 1 + \frac{87Sr}{^{87}Rb}$$

A useful method for determining the primary ^{87}Sr/^{86}Sr ratio in a rock is by constructing a Rb-Sr isochron. This is done by analyzing several minerals and a whole-rock sample from the same rock body to give several data points for plotting and calculations. If the same initial ^{87}Sr/^{86}Sr was trapped in the different phases and they have remained as closed systems over the age of the rock then the minerals with higher Rb contents will have higher ^{87}Sr/^{86}Sr ratios and a plot of total present-day ^{87}Sr/^{86}Sr vs. ^{87}Rb/^{86}Sr will be a straight line. The initial ^{87}Sr/^{86}Sr ratio will be defined by the intercept with the zero ^{87}Rb/^{86}Sr ratio and the time of formation of the rock will be given by

$$t = \frac{1}{\lambda} \ln (1 + slope)$$

An interesting Rb-Sr isochron or strontium evolution diagram for rocks from the Godthaab area of Greenland is shown in Figure 7.12. The rocks analyzed appear to be part of a uniform system, fall on a straight line, and give a primordial ^{87}Sr/^{86}Sr value of 0.7009. The slope yields an age of about 3740 million years.

Isochrons have proven to be very useful in the study of rocks that give different or discordant ages in an area. Studies show that whole-rock samples usually less than a cubic meter in size are closed systems during metamorphism while mineral grains in the rocks may experience exchange and homogenization of the Sr isotopes during the heating event.

Such an homogenization causes a resetting of the isochron to a horizontal line as illustrated in Figure 7.13. At this time the radioactive clock would be reset and the strontium would develop a new isochron to give the date of the metamorphic event. However, the isochron through the whole-rock analyses would not be altered, and the analyses and plotting of different whole-rock samples would provide a date for the original event. A diagram illustrating an ideal two-stage isochron for both rocks and minerals indicating both the initial igneous and later metamorphic events is shown in Figure 7.14. In practice some areas have such a complicated geologic history that the strontium evolution diagram is difficult if not impossible to interpret.

The determination of initial ^{87}Sr/^{86}Sr ratios can provide interesting information about the origin of a rock. If it is assumed that most of the earth's Rb is concentrated in the crust, magmas derived from the crust will have higher ^{87}Sr/^{86}Sr ratios than those derived from the mantle. Such an assumption is supported by the fact that oceanic basalts have initial ^{87}Sr/^{86}Sr values of about 0.702 to 0.705, leading to the conclusion that the Rb/Sr value in the upper mantle is about 0.02. Initial ratios in
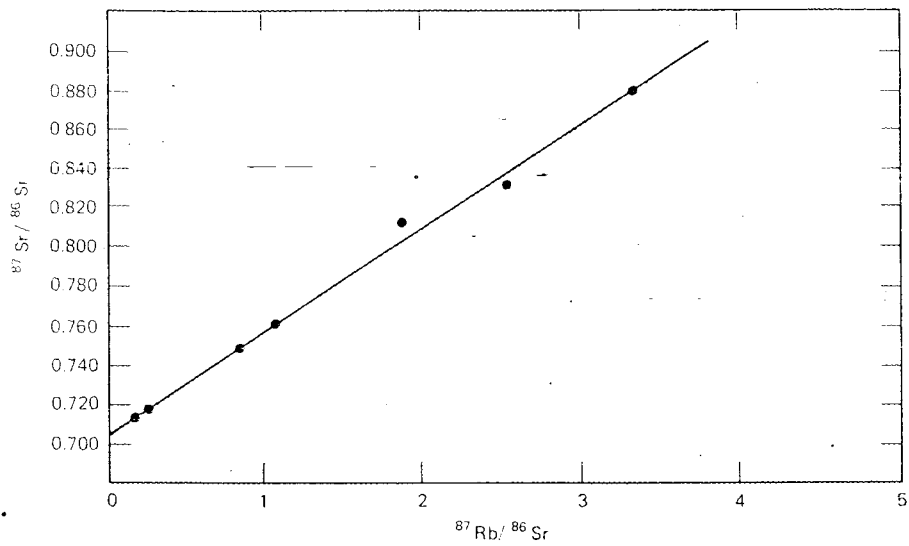
Figure 7.12 Isochron diagram for stony meteorites. Achondrites fall into the open circle at the left, and chondrites are solid dots.
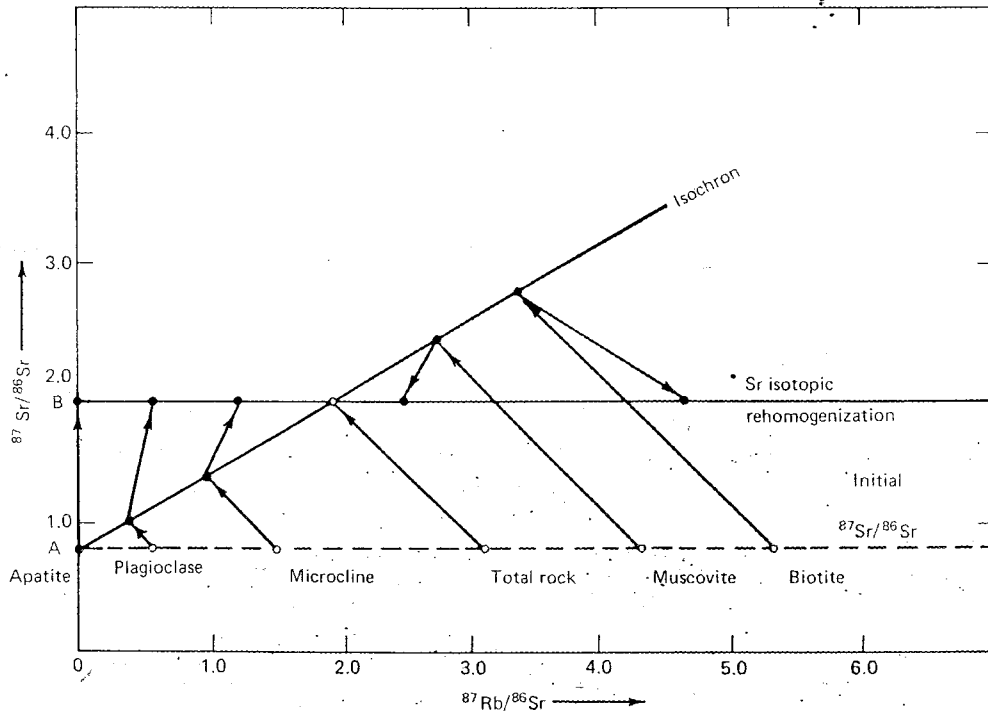
Figure 7.13 Isochron diagram showing the effect of episodic homogenization of Sr. The total rock and component minerals had an initial Sr^{87}/Sr^{86} indicated by A. Later, a metamorphism caused the Sr composition in the minerals to rehomogenize to the total rock ratio (B) existing at that time.
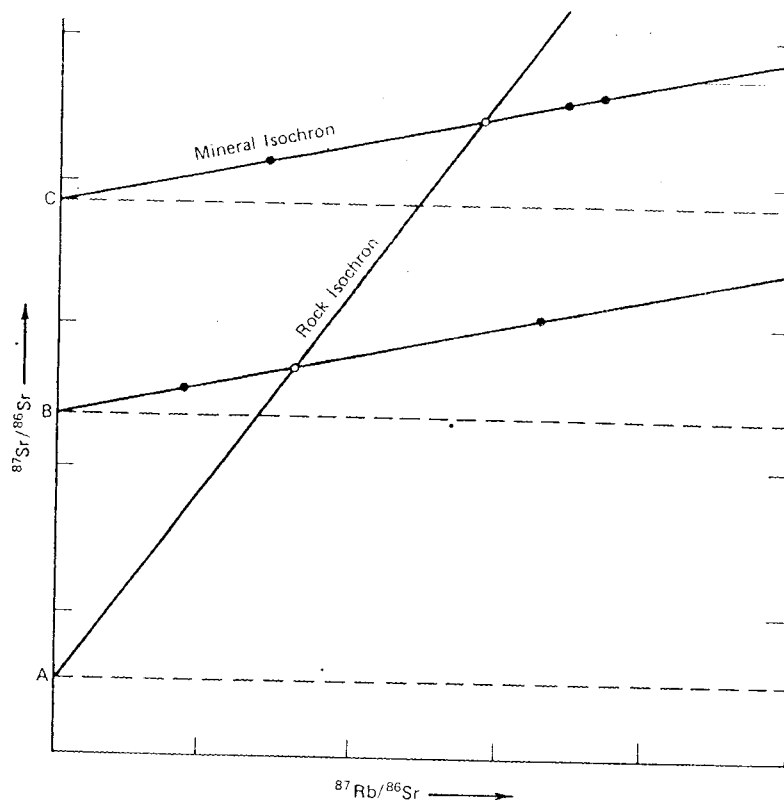
Figure 7.14   Schematic isochron diagram for whole rocks (open circles) from the same geologic system, and minerals (solid circles) comprising the whole rocks. The whole-rock isochron corresponds to the primary age, the mineral isochrons to the age of metamorphism. The initial $^{87}Sr/^{86}Sr$ for the whole rock is A, and $^{87}Sr/^{86}Sr$ in individual whole-rock specimens at the later time of metamorphism is denoted by B and C.

granitic igneous rocks may be in the range of 0.704 to 0.711 with Rb/Sr ratios of about 0.25. Magmas derived from melted crustal rocks will thus have higher $^{87}Sr/^{86}Sr$ ratios than those from an upper mantle source.

Carbon Dating

The carbon-14 method used for dating relatively recent geologic events is a decay clock rather than an accumulation clock. Unlike the radioactive isotopes of U, Th, K, and Rb, which have survived since the origin of the elements, carbon-14 is being continually produced by the capture of neutrons by nitrogen-14 in the upper atmosphere. The earth is constantly being bombarded by cosmic rays that produce nuclear particles upon collision with atmospheric molecules. Neutrons are captured by nitrogen according to the equation

$$^{14}N + n \rightarrow {}^{14}C + H$$

The ^{14}C is radioactive with a half-life of about 5730 years. The carbon formed reacts quickly with oxygen to form radioactive CO_2 which enters the photosynthetic cycle together with stable carbon CO_2. The partially radioactive atmospheric CO_2 enters and exchanges rapidly with the hydrosphere and biosphere with respect to the ^{14}C half-life. Living plants, animals, and their carbonate hardparts had a natural radioactivity of about 15 disintegrations/min-g of carbon until it was changed by the activities of humans. The burning of fossil fuels has added CO_2 to the air to provide a dilution of about 2% of the normal value. The dilution of ^{14}C by this extra CO_2 has been more than counterbalanced by the production of ^{14}C by thermonuclear explosions. Bomb testing has temporarily increased the ^{14}C to almost double the normal level.

The carbon-14 decay clock is started when a piece of wood or other material stops its exchange with the atmospheric reservoir by death. When this occurs, decay of ^{14}C sets in and allows the time of the start of the clock to be estimated. The precision of this calculated date is controlled by analytical technique, contamination, and the nature of the sample. A piece of trunk wood for example has a range of initial dates depending on the location of the sample with respect to the core of the trunk. A radioactive date may be calculated from the equation

$$t = \frac{1}{\lambda} \ln \frac{Ne}{Np}$$

where Ne is the equilibrium ^{14}C activity and Np is the present ^{14}C activity. The activity is most often determined by converting the cleaned carbon samples to CO_2 and counting the CO_2 in low-background proportional counters. By convention the ages are calculated using a decay half-life of 5568 years, which is about 3% less than the recently determined half-life values. If a large enough sample is available it is possible to obtain a date (with a maximum error of about ± 150 years) of a sample up to 50,000 years old.

Analytical factors limit the range of ^{14}C dating to about 50,000 years. The interpretation of ^{14}C has been invaluable in elucidating the geologic events of the recent Pleistocene and Holocene, and in providing a solid time foundation for archaeological specimens.

Care must be taken in evaluating and using ^{14}C dates. It is possible for old or ^{14}C free carbon to be added to some geologic reservoirs such as when deep ocean waters are carried to the surface or limestone dissolves in lakes. Basic to the assumptions upon which ^{14}C dating is founded is an assumed, relatively constant rate of ^{14}C production. Careful measurements of tree-ring dated woods shows this is not exactly true. If necessary, corrections may be attempted for such variations.

SELECTED REFERENCES

Craig, H. (1961). Standard for reporting concentrations of deuterium and oxygen-18 in natural waters. *Science* **133**, 1833.

Faure, G. (1977). *Principles of isotope geology*. 464 pp. John Wiley & Sons, New

York. Covers both stable and radioactive isotopes in detail. Many useful references.

Fleischer, R. L., P. Price, and R. M. Walker (1975). *Nuclear tracks in solids: Principles and applications.* 605 pp. University of California, Berkeley. A summary of work on this interesting technique.

Harper, C. T. (ed.) (1973). *Geochronology-radiometric dating of rocks and minerals.* 469 pp. Dowden, Hutchison & Ross, Stroudsburg, Pa. A volume in the Benchmark Papers in Geology. Reprints, classical papers in geochronology.

Hoefs, J. (1973). *Stable isotope geochemistry.* 140 pp. Springer-Verlag, New York. A very useful review of stable isotopes. Contains extensive references to earlier work.

O'Nions, R. D., S. R. Carter, N. M. Evensen, and P. J. Hamilton (1979). Geochemical and cosmochemical applications of Nd isotope analysis. *Ann. Rev. Earth Planet. Sci.* 7, 11–38.

Ralph, E. K., and H. N. Michael (1974). Twenty-five years of radiocarbon dating. *Amer. Sci.* 62, 553–560. The history of this useful technique.

Taylor, H. P., Jr. (1974). The application of oxygen and hydrogen isotope studies to problems of hydrothermal alteration and ore deposition. *Econ. Geol.* 69, 843–883. One of a series of excellent papers by this author on stable isotopes. Extensive bibliography references important earlier work.

Urey, H. C., H. A. Lowenstam, S. Epstein, and C. R. McKinney (1951). Measurement of paleotemperatures and temperatures of the Upper Cretaceous of England, Denmark and the Southeastern United States. *Bull. Geol. Soc. Amer.* 62, 399. Defines the standard for oxygen isotope analysis.

York, D. and R. M. Farquhar (1972). *The earth's age and geochronology.* 178 pp. Pergamon Press, Oxford. A well-balanced introduction to age dating techniques.

8 THE ATMOSPHERE

INTRODUCTION

Although the atmosphere makes up only a small amount of the total mass of the earth, its influence on the chemistry of the earth's surface is appreciable. It acts as a transporting medium for the components of many of the geochemical reactions that take place on the surface of the earth, such as the transport of oxygen, carbon dioxide, and water, which are essential in the chemical weathering process. The atmosphere is an important part of the geochemical cycle of many elements including sulfur, carbon, and the halogens derived from sea spray or biochemical processes and distributed by atmospheric turbulence. Some minor constituents of the atmosphere are geochemically important. Carbon dioxide, whose concentration is only a few hundredths of one percent, provides the raw material for plant life. Ozone, most of which is present in a diffuse layer (the ozonosphere) within the stratosphere, plays a vital part in absorbing ultraviolet radiation; if this adsorption did not take place the ultraviolet radiation at the earth's surface would be fatal for most forms of life. The presence of an atmosphere is essential for a liquid phase or hydrosphere to exist.

THE COMPOSITION OF THE ATMOSPHERE

The atmosphere is simple in composition, being made up almost entirely of three elements: nitrogen, oxygen, and argon. Some of the components, including the major constituents, are constant in concentration at least over short-term variations. Although we have evidence that some have changed slowly over geological time, the geologic record indicates the atmosphere has had a fairly constant composition over

210

Table 8.1 The Composition of the Atmosphere at Ground Level

Gas	Concentration by Volume (ppm)	Concentration by Weight (ppm)
Nitrogen, N_2	780,840	755,220
Oxygen, O_2	209,460	231,400
Argon, Ar	9340	12,880
Carbon Dioxide, CO_2	320	490
Neon, Ne	18.18	12.67
Helium, He	5.24	0.724
Krypton, Kr	1.14	3.30
Xenon, Xe	0.087	0.39
Hydrogen, H_2	0.55	0.038
Nitrous Oxide, N_2O	0.33	0.50

the last 10^8 to 10^9 years. Other components, notably water, show great variations over short periods of time. Water vapor in the atmosphere may range up to to 4%. The carbon dioxide concentration in the atmosphere is being increased with the increased consumption of fossil fuels. The average composition of the atmosphere at ground level is given in Table 8.1. According to Avogadro's law, the volume concentrations give directly the relative number of the different molecules and atoms.

Convection currents maintain constancy of the proportions of the different components at elevations below 60 kilometers. Above that level, gravitational separation according to molecular weight may be observed. For this reason, the layer below 60 kilometers is often spoken of as the homosphere. The geochemistry of the atmosphere concerns itself mainly with this region and its reactions with the lithosphere, biosphere, and hydrosphere. Geophysicists, however, have studied the upper atmosphere extensively and have shown that many interesting physical and photochemical reactions take place in the region above 60 to 100 kilometers and as the atmosphere diffuses into space. Figure 8.1 illustrates some of the properties, structure, and nomenclature of the atmosphere. The atmosphere gradually thins out into interplanetary space, and above an altitude of 600 kilometers the atmospheric molecules describe free elliptical orbits in the earth's gravitational field. Figure 8.2 shows the relationship of the constituents to each other and the changes that take place as a function of altitude. The atmospheric components are given as number density or the number of molecules per cm^{-3}. Note the parallel lines at the lower altitudes for most constituents. Notable exceptions are H_2O, O_3, and atomic oxygen. Atomic oxygen produced by photochemical reactions appears at lower altitudes than atomic nitrogen because of the greater bond strength of the N_2 molecule. It may be seen that at an altitude of approximately 240 kilometers atomic oxygen becomes the predominant atmospheric species. At higher elevations He and H predominate.

Although in composition the atmosphere can be looked upon as becoming uniformly more diffuse upward, it does show some structural features of remarkable physical significance. The lower part of the atmosphere in which convection is prominent is known as the troposphere. Above this is the stratosphere, so named because

·it appears to be stratified in a number of layers between which no strong vertical circulation seems to exist. The boundary between the troposphere and the stratosphere varies both with the latitude and with the season, but can be placed at a height of about 10 to 15 kilometers. In the troposphere, the temperature decreases with increasing height, whereas in the stratosphere the temperature is independent of altitude. Above about 80 kilometers, the stratosphere passes into the ionosphere. In this region, the atmosphere is rendered conducting by ionization induced by ultraviolet radiation from the sun.

The calculated mass of the atmosphere is about 50×10^{20} g; one of the latest and most careful estimates gives 51.17×10^{20} g for the total mass of dry atmosphere, that is, without any water vapor. In considering the chemical composition of the atmosphere, it is important to remember that the composition even for the nonvariable elements is not a static equilibrium but is kept constant by balancing kinetic processes. An important concept in considering these mobile or fugitive elements that make up the atmosphere is the term "mean lifetime" which gives the mean, or average, time that a given molecule resides in the atmosphere. For oxygen it appears that the average oxygen molecule is recycled about every 2000 years. For carbon dioxide, the mean residence time is approximately 30 years. For nitrogen, the time is much
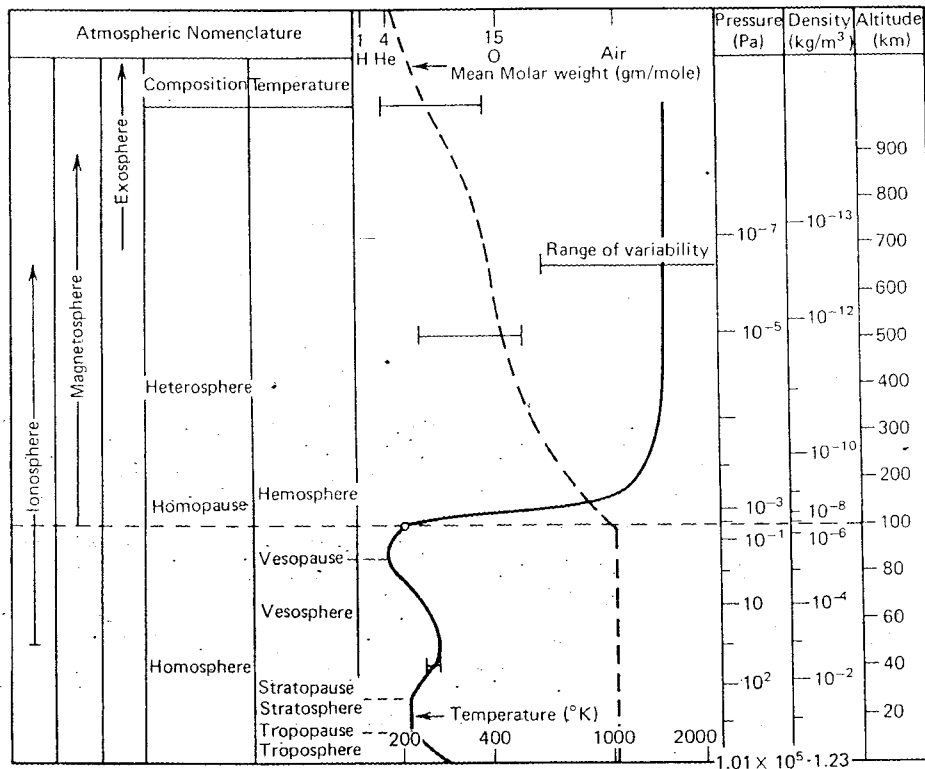
Figure 8.1   Structure and composition of the atmosphere.
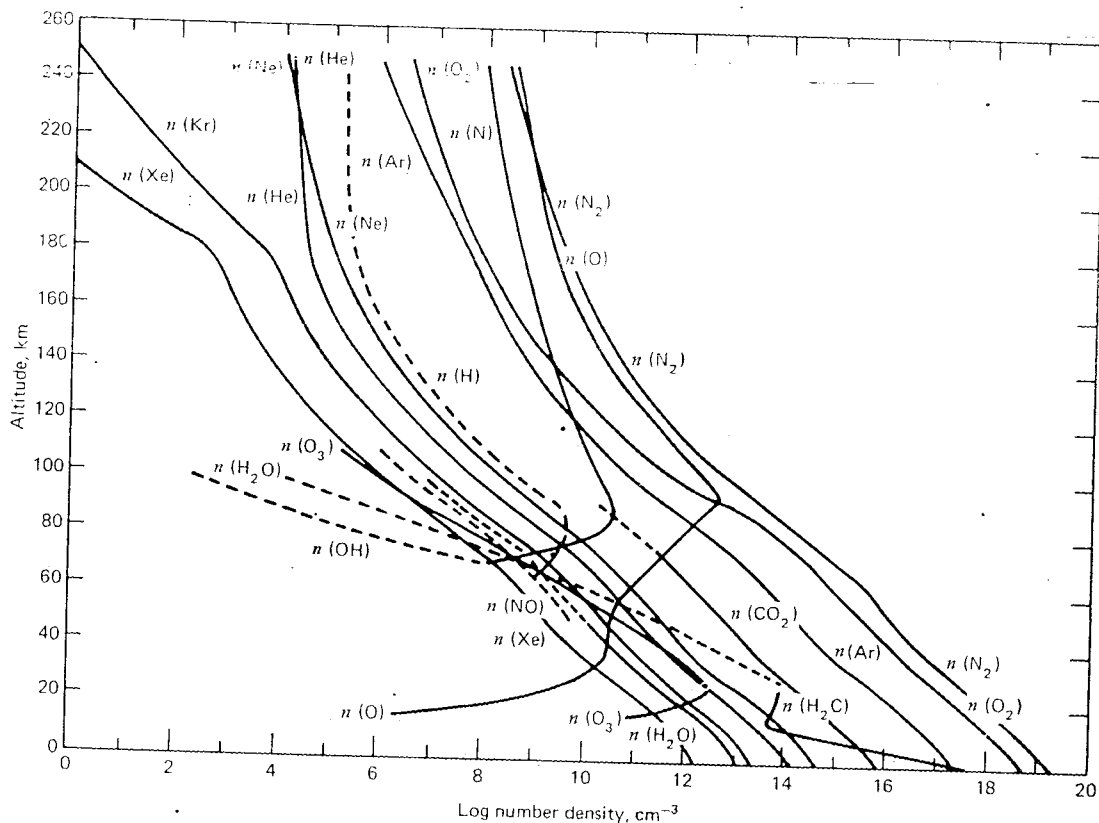
Figure 8.2 Vertical distribution of atmospheric constituents from surface to 250 km. (After Air Force Geophysics Research Directorate, 1960.)

longer because the rate of exchange is small compared to the total mass of nitrogen in the atmosphere. The calculations indicate that the average nitrogen molecule should reside in the atmosphere from 10 to 100 million years before it is replaced. Feedback mechanisms exist that control the overall stability of the major atmospheric constituents.

Figure 8.3 on page 216 shows a schematic balance sheet for nitrogen. This major constituent of the atmosphere is primarily found in the molecular form as N_2 but does enter into a number of reactions to provide for a cyclical transfer of individual atoms to and from the atmosphere. A similar diagram for carbon is included in Chapter 10. Note the predominating reactions controlling the movement of nitrogen are those of biological origin. Photochemical and lightning discharge reactions also produce a noticeable transfer of molecular nitrogen to nitrate nitrogen. Biological fixation of nitrogen may involve either oxidation to produce nitrate and nitrite or reduction to produce ammonia. These forms may be then utilized by bacteria to provide them with a source of energy. Molecular nitrogen is returned to the atmosphere in the energy release processes.

The balance of oxygen in the atmosphere is regulated by the biosphere through photosynthesis. The exchange rate is much greater than that of nitrogen so the oxygen is passed through the cycle in about 2000 years. The lack of evidence for major short-term fluctuations in the oxygen content of the atmosphere in recent geologic time suggests that the addition and depletion processes must be well balanced. The main source of oxygen is photosynthesis, while the main sinks are respiration and oxidation by decay or weathering. About 99% of the oxygen released by photosynthesis is used up by oxidation of the organic material produced in the same process. The excess oxygen is available because of burial of organic matter before it can be reoxidized to CO_2. Since there is evidence that the oxygen content of the atmosphere has increased over geologic time, it is interesting to speculate whether oxygen has reached a steady-state plateau or will continue to slowly increase in the future.

Carbon dioxide, like oxygen, is involved in photosynthesis but the main process controlling its concentration in the atmosphere is its solubility and reaction with water. Carbon dioxide is much more soluble in water than oxygen and nitrogen. The proportional solubility of CO_2, O_2, and N_2 is $70:2:1$. Most free oxygen and nitrogen are in the atmosphere while most free carbon dioxide is in the oceans. In addition much carbon dioxide is also tied up as calcium carbonate on the ocean floor or as ancient carbonate sediments on land. Carbon dioxide in the sediment-ocean-atmosphere system buffers its content in the atmosphere to an essentially constant value.

· Carbon dioxide's role in photosynthesis may be generalized by the equation

$$nCO_2 + nH_2O = (CH_2O)_n + nO_2$$
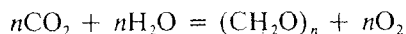
The rate of the exchange of carbon dioxide is rapid through the photosynthetic process and the reverse process of respiration or decay.

In addition to carbon dioxide, carbon also is found in the atmosphere in the form of carbon monoxide, methane, and more complex organic molecules. The carbon monoxide may be produced by photochemical oxidation of organic matter or by dissociation of CO_2. The methane is primarily produced by bacterial decay of organic matter. It is removed from the air by oxidation.

As indicated in Table 8.1 the other measured constituents of the atmosphere are the noble gases, notably argon. They are atomophilic elements since they do not enter into chemical reactions that can cause them to be removed. All are present in small amounts, smaller than their molecular weights in relation to nitrogen might lead us to expect. Argon is anomalously abundant with respect to the other noble gases because it is a daughter product formed by the radioactive decay of ^{40}K in the earth's crust and has been accumulating steadily over geologic time. Helium, a radioactive decay product as alpha particles, is depleted because of its chemically inert nature and low atomic weight which allows it to be slowly lost by diffusion into space.

Variable Constituents of the Atmosphere

So far we have been discussing the constituents of the atmosphere that are present in all samples of the air in amounts that do not vary appreciably from time to time or from place to place. Other constituents must also be taken into account. Variable

Table 8.2   Variable Constituents of the Atmosphere at Ground Level

Gas	Concentration (ppm by Volume)
H_2O	200 – 40,000
O_3	0.07
SO_2	<0.001 to 1
H_2S	to 0.0002
CO	to 0.2
CH_4	to 2
N_2O	to 0.02
NH_3	to 0.02

constituents of the atmosphere are given in Table 8.2. Of these, the most important is water vapor, which may be present in amounts varying from 0.02 to 4% by weight. The content of water vapor depends on a number of factors, of which temperature is the most significant; thus the average annual values for different latitudes are 0°, 2.63 vol.,%; 50° N, 0.92 vol. %; 70° N, 0.22 vol. %. Water vapor in the atmosphere plays an important part in regulating climatic conditions. Water vapor, together with carbon dioxide, is responsible for the important temperature controlling mechanism known as the greenhouse effect. Solar radiation of certain wavelengths may pass through the H_2O and CO_2 molecules and strike the earth's surface. Here the radiant energy is reemitted at longer wavelengths and is absorbed by the H_2O and CO_2 thereby warming up the atmosphere. It has been estimated that if the earth were 5% closer to the sun the increased radiation would be enough to produce a runaway greenhouse effect which would volatilize the earth's hydrosphere and produce quite a different planet from the one on which we now live. On the earth today the absorption of heat brings about a closer approach to uniformity in the temperatures of different latitudes than would exist on a dry earth.

Sulfur compounds are present in the atmosphere in variable quantities and may be considered as contaminants rather than normal constituents. Hydrogen sulfide is a product of putrefaction, and small quantities are also given off, together with sulfur dioxide, by volcanoes. An important local source of sulfur compounds in the air is the combustion of coal, and the highest concentrations of such compounds are found in industrial areas. Sulfurous gases are temporary constituents of air insofar as they are readily soluble in water and are washed out by rain. The total amount of sulfur thereby circulated through the atmosphere is by no means insignificant; careful measurements over a period of years at Rothamsted, England, showed that the equivalent of 7.8 kg of SO_3 was annually precipitated on each acre of land at the agriculture experimental station there. In addition the atmosphere transports sulfate sulfur derived from sea salts by the same process as the sodium chloride discussed below. This material may be deposited in the rain or by the dry-fall of aerosols from the atmosphere. It has been estimated that about 30% of atmospheric SO_2 results from human activities. The average global residence time for SO_2 is about 40 days.

Measurements have been made in many localities to evaluate the amount of combined nitrogen brought to the surface of the earth in solution in rain water. These figures vary from 1 to 6 kg per acre per year, mainly in the form of ammonium nitrate or ammonium sulfate. A large part of this has been added to the atmosphere by the decomposition of organic matter at the earth's surface, but some is due to the inorganic production of oxides of nitrogen. Nitrous oxide is produced in the soil by bacterial activity, and the small concentration of this gas in the atmosphere can be regarded as a steady state between the biological production and the direct or indirect decomposition of the gas in the upper atmosphere. The maximum nonbiological fixation of nitrogen is estimated to be about 0.0035 mg/cm^2 of surface per year, corresponding to about one-seventh of the mean annual precipitation of nitrate nitrogen in temperate localities. Figure 8.4 illustrates a detailed balance sheet for all the compounds of nitrogen between the atmosphere, the biosphere, and the ocean. Although the figures given are estimates from many sources, it may be seen that there are many small, complex, and rapid cycles involved in the distribution of nitrogen species in the atmosphere.

Of the temporary constituents of the atmosphere, sodium chloride is very important. Figures for the amount of this substance precipitated upon the land by rain and aerosol dry-fall range from 9 to 90 kg per acre annually. Most if not all represents salt raised by spray from the ocean. As would be expected, sodium chloride in the air is greatest near the sea and rapidly diminishes away from the coast. The atmospheric circulation of salt has received much attention, since cyclic salt of this kind is largely responsible for the sodium chloride of inland waters.

The other halogens—fluorine, bromine, and iodine—have also been detected in the atmosphere. Fluorine is probably put into the atmosphere as an industrial contaminant liberated by the burning of fuel and the calcination of material containing fluorine, or released in the manufacture of phosphate fertilizers from rock phosphate, which always contains some fluorine-bearing apatite. In places this results in an
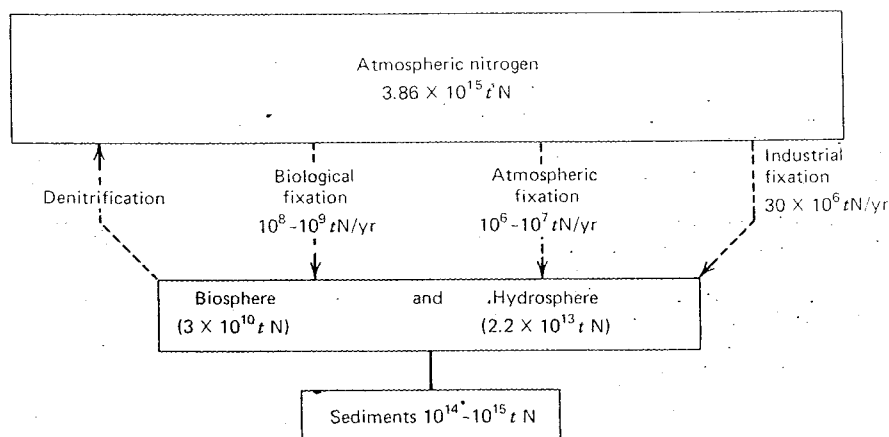
Figure 8.3 Geochemical cycle of nitrogen.
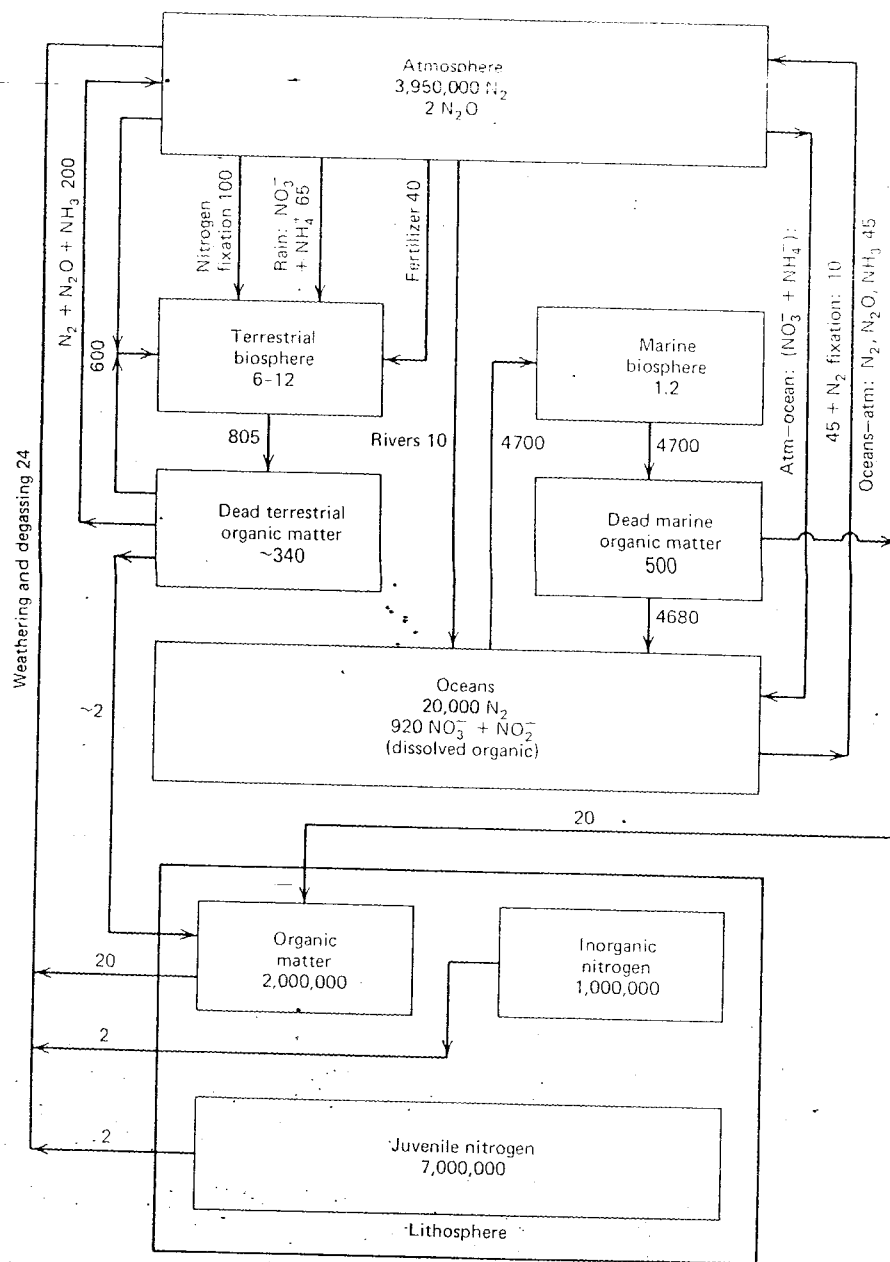
Figure 8.4 Major aspects of the nitrogen cycle; nitrogen in reservoirs in units of 10^{15} g; transfer rates in units of 10^{12} g/yr. (Holland, 1978, p. 306)

accumulation on grasslands that may be harmful to animals. Fluorine as HF has been released by volcanic activity in Iceland, causing a serious problem with fluorine poisoning of sheep after volcanic eruptions. Bromine and iodine appear to be universally present in the atmosphere in very small and variable amounts; analyses of rain water show 0.03 to 0.002 mg/liter of bromine and 0.002 to 0.0002 mg/liter of iodine. The ocean is probably the main source of these elements, although some are contributed by industrial gases. The ratio of Cl:Br:I is 100:0.34:0.00021 in sea water and about 100:5:0.5 in rain water; this indicates a moderate enrichment of bromine and an enormous enrichment of iodine in atmospheric precipitation. Iodine is readily liberated from its compounds, and because it has an appreciable vapor pressure at ordinary temperatures it can easily exist free in the atmosphere.

The influence of the internal combustion engine on atmospheric geochemistry has become increasingly marked in recent years, especially in large cities. The automobile is a primary source of carbon monoxide; atmospheric carbon monoxide begins to be hazardous to humans at concentrations of about 100 ppm, if this level is maintained for several hours. Such concentrations may be reached occasionally in areas of heavy traffic. When combustion takes place at high pressures, as in the cylinders of an internal combustion engine, poisonous nitric oxide is produced, and this is oxidized in the atmosphere to nitrogen dioxide, a still more toxic gas. A further source of atmospheric pollution is the lead compounds used as antiknock agents in gasolines. Since leaded gasolines were introduced in 1923 more than 2.6 million tons of lead have been converted into lead alkyls and burned in automobiles. The present annual production of lead alkyls is equivalent to about 300,000 tons of lead. The average cumulative lead contamination in the northern hemisphere is about 10 mg/m^2 from gasoline burning alone. In highly industrialized and motorized areas the amount of lead voided into the atmosphere is many times greater than the average.

THE EVOLUTION OF THE ATMOSPHERE

The evolution of the atmosphere is of great interest and has been the object of extensive study. The question is primarily one of long-term cycles or changes of the atmospheric components as opposed to the shorter term balance discussed in the previous sections. The investigation of these cycles and changes is difficult since samples of early atmospheres are not available. However there may be clues contained in ancient sediments formed in contact with these atmospheres and in evidence related to the origin and changes in living organisms. As we have seen, plants play a critical role, particularly in the control of free oxygen in the atmosphere, and the history of oxygen apparently is intimately tied to the evolution of plant life. The evolution of the atmosphere can be discussed in terms of the following:

1. The composition of the primeval atmospheres.
2. Additions during geological time.
3. Losses during geological time.

Although the quantitative results produced by these processes are often difficult to evaluate, considerable data have been accumulated. These data are discussed briefly in the following sections.

The Composition of the Primeval Atmosphere

Our ideas as to the composition of the primeval atmosphere are conditioned largely by the mode of origin that we ascribe to the earth and to its very early history. According to the planetesimal hypothesis, the particles aggregating to form the earth had no atmosphere associated with them; the primeval atmosphere originated from the gases occluded or combined within the planetesimals and released by the heat and chemical reactions accompanying and following aggregation. On this basis the constituents of the atmosphere have been largely, if not entirely, exhaled from the interior of the earth. The strongest evidence for this lies in a comparison between terrestrial and cosmic abundances of the elements. Table 8.3 provides such a comparison in terms of a deficiency factor, based on the assumption that silicon, as an element forming stable, nonvolatile compounds, has the same relative abundance both in the earth and the solar system. In the last column of Table 8.3 the deficiency factor is expressed on a logarithmic scale in which zero represents comparable abundance in the earth and throughout the solar system, that is, neither enrichment nor deficiency. The truth of the assumption that silicon is an element that has the same relative abundance in the earth and in the solar system is supported by the fact that comparable elements, such as sodium, magnesium and aluminum also have deficiency factors near zero; that is, their terrestrial and cosmic abundances are of the same order. The degree of precision of these deficiency factors is not great, but qualitatively they are undoubtedly significant. The much greater deficiency factor of helium than hydrogen can be ascribed to the retention of hydrogen by chemical combination with other elements; helium, being an inert gas, cannot be retained in this

Table 8.3 The Abundances of Some Elements in the Earth and in the Solar System as a Whole

	Atomic Number	(atoms/10,000 atoms Si)		Deficiency Factor $\log(b/a)$
		Whole Earth (a)	Solar System (b)	
H	1	84	2.2×10^8	6.4
He	2	3.5×10^{-7}	1.4×10^7	15
C	6	71	94,000	3.1
N	7	0.21	29,000	5.1
O	8	35,000	280,000	0.9
F	9	2.7	8	0.5
Ne	10	1.2×10^{-6}	8,300	9.8
Na	11	460	600	0.1
Mg	12	8,900	10,600	0.1
Al	13	940	850	0
Si	14	10,000	10,000	0
P	15	100	130	0.1
S	16	1,000	5,020	0.7
Cl	17	32	57	0.3
Ar	18	5.9×10^{-4}	220	5.6
Kr	36	6×10^{-8}	0.87	7.2
Xe	54	5×10^{-9}	0.015	6.5

Table 8.4   Summary of Data on the Probable Chemical
              Composition of the Atmosphere during Stages 1, 2,
              and 3

	Stage 1	Stage 2	Stage 3
Major components $P > 10^{-2}$ atm	CH_4 H_2 (?)	N_2	N_2 O_2
Minor components $10^{-2} < P < 10^{-4}$ atm	H_2 (?) H_2O N_2 H_2S NH_3 Ar	H_2O CO_2 Ar	Ar H_2O CO_2
Trace components $10^{-4} < P < 10^{-6}$ atm	He	Ne He CH_4 NH_3 (?) SO_2 (?) H_2S (?)	Ne He CH_4 Kr

way. A similar relation is evident when the deficiency factors of oxygen, nitrogen, and neon are compared; clearly, chemical activity has played a significant role in the retention of the gaseous elements. The deficiency of these lighter constituents in the earth is explicable with either the nebular or the planetesimal hypothesis; however, as Table 8.3 shows, the heavy inert gases are also deficient. Krypton and xenon are about a million times less abundant in the earth than their immediate neighbors in. the periodic table. Since their nuclear properties are not strikingly dissimilar to those of their neighbors, it is reasonable to expect these rare gases to be of the same order of abundance. If the earth condensed directly from solar matter, krypton and xenon, unlike the lighter elements, should have been completely retained because of their high atomic weight. Their deficiency therefore is an important point in favor of some form of the planetesimal hypothesis whereby some fractionation of volatile elements took place before the earth's accretion.

The evolution of the atmosphere on an earth formed by planetesimal accretion has been carefully discussed by Holland (1962). He distinguished three stages in its evolution (Table 8.4). The accretion of the earth was an exothermic process, and this, plus radiogenic heat, would soon result in magma generation. Holland speculates that the chemistry of volcanic gases at the first stage in the evolution of the atmosphere was probably very different than it is today. Before the formation of the core, metallic iron would be present in the upper part of the mantle (if the planetesimals were similar in composition to ordinary chondrites, the metal would be part of the accreting material; if they were like carbonaceous chondrites, metallic nickel-iron would be formed as the material heated up). Volcanic gases in equilibrium with free iron would consist largely of H_2, H_2O, and CO, with minor amounts of N_2, CO_2, and H_2S. Hydrogen was probably the dominant constituent. As the gases cooled the CO and CO_2 reacted with H_2 to form CH_4. Whether nitrogen may have reacted to form NH_3 depends on whether the rate of escape of hydrogen from the planet was suffi-

ciently slow to permit the existence of an appreciable hydrogen pressure. If the earth's surface was cool enough and nearly all the water condensed, the main constituents of the atmosphere would then be CH_4 and H_2.

The duration of the first stage was determined by the time that elapsed between the accretion of the earth and the formation of a core by the gravitational separation of the metallic iron from the mantle. Holland suggests that this did not exceed 500 million years. The removal of the metallic iron phase left a siliceous mantle essentially like that existing today, and subsequent volcanism was accompanied by gases much more oxidized than those of the first stage, and probably similar in composition to the Hawaiian volcanic gases today. Water was the dominant component, and CO_2, CO, H_2, SO_2, and N_2 minor constituents. The atmosphere during the second stage would contain largely N_2 with minor amounts of CO_2 and H_2O. Photochemical dissociation of water vapor in the upper atmosphere would produce oxygen and hydrogen, with the hydrogen escaping into outer space. However, the free oxygen would not accumulate in the atmosphere, but would be used up in oxidizing the more reduced constituents of the volcanic gases.

The second stage came to an end when oxygen production exceeded oxygen use. The problem of the time of appearance of free oxygen and its mode of formation is an intriguing one. Increasing knowledge and understanding of Precambrian rocks has contributed greatly to the solution of this problem. Cloud (1976) has summarized the combined evidence of paleontology, stratigraphy, and geochemistry. The evidence indicates that green-plant photosynthesis existed at least 1.7 to 2.0 billion years ago and that atmospheric oxygen first began to be available in relatively large quantities probably about 1.2 billion years ago. From the time of the first available photosynthetic oxygen until about 1.2 billion years ago ferrous iron, carried into large water bodies or the ocean by reducing surface waters, may have served as a vast oxygen sump, retarding the evolution of free oxygen to the atmosphere. Cloud remarks that the existence and facies of the unique sedimentary Precambrian iron formations of the Lake Superior type possibly reflect the onset and nature of early photosynthesis. Typical Precambrian iron formation rocks range in age from about 2.7 to 1.7 billion years before the present. The oldest fossils, of possibly algal origin, are apparently somewhat older than 2.6 billion years. Cloud suggests that the appearance of oxygen-generating organisms in the hydrosphere was necessary to trigger the deposition of this type of iron formation, and that deposition ceased at some point of compensation or evolution between 1.7 and 1.2 billion years ago. As early as about 1.2 billion years ago there was enough oxygen in the atmosphere to produce extensive red beds (nonmarine sedimentary rocks colored by ferric oxides). By the beginning of Paleozoic time, about 600 million years ago, enough oxygen had accumulated in the atmosphere to permit the evolution of the metazoa (multicelled animals). Thus the buildup of oxygen in the atmosphere was gradual and did not reach its present concentration until quite late in earth history.

Circumstantial support for a reducing atmosphere during the early history of the earth also comes from scientists interested in the origin of life. Many theories, following the early work of Miller and Urey (1956) on the production of organic molecules by energy release in simulated protoatmospheres, are based on the predicted presence of CH_4, H_2O, and sometimes NH_3. Mixtures proposed are often based on

analogy of the reduced atmospheres observed on the major planets. In many respects such theories are difficult to reconcile with the low noble gas abundances, which indicate that the earth was devoid of an atmosphere sometime during its early history, whereas the atmospheres of the Jovian planets appear to be primary and captured during each planet's accretion. The earth may have lost its primeval atmosphere during a hot early history or by an explosive stage in the development of the sun known as the Hitashi stage, during which any truly primitive atmosphere on the earth would be blown away.

To understand the chemistry and evolution of the atmosphere of the earth, it is also important to study the atmospheres of the other planets. Such a study has often been termed "comparative planetology." Since the atmospheres of the other planets are accessible to observation either directly or with space probes, their composition is among the first planetary data that may add important information to enable us to better understand the chemistry of the earth's atmosphere.

We have seen that the planets can be divided into the terrestrial planets similar to the earth and the large outer planets, which have different properties, primarily in their atmospheres, due to their large size and low temperatures.

For comparison with the earth's atmosphere the two most important planets are Venus and Mars. We have seen that Venus may be considered a twin of the earth in terms of its size. Likewise, we might expect that the atmospheric evolution may show some similarities. Mars, being smaller, has lost gases by diffusion at a greater rate than the earth. The measured atmospheric pressure on Mars is 7.6 mbar. Direct measurement of Mars' atmospheric composition and isotopic ratios allow comparisons with the earth's atmosphere (Table 8.5). Martian atmospheric carbon and oxygen isotope ratios are similar to the earth's. The ratio $^{14}N/^{15}N$ is 277 for the Earth and 165 for Mars, while $^{40}Ar/^{36}Ar$ is 292 for the Earth and 3000 for Mars. The present atmosphere on Mars represents a small part of the total atmosphere evolved since its formation. The relative abundances of noble gases show a similarity to those on the earth and in ordinary chondritic meteorites. The atmospheric evolution of the atmospheres of the two planets is significantly different. They may have accreted different proportions of volatiles during formation and the smaller mass of Mars has permitted greater diffusion of gas away from the planet.

Observations of the atmosphere of Venus indicate that it is made up primarily of carbon dioxide. This has been determined both by spectroscopic measurements from Earth-based telescopes and fly-by vehicles, and by a direct measurement by a Russian Venus landing. Indirect methods make it difficult to determine the nitrogen content because the spectroscopic detection methods are not sensitive to this element. The atmosphere of Venus has been reported to be from 10 to 100 times more dense than Earth's. The composition is approximately 95% CO_2, 1% H_2O, 2% N_2, and less than 0.01% O_2. The surface temperatures of Venus are close to $600\,^\circ C$. With such temperatures, it is impossible that the surface of Venus would have a hydrosphere and, hence, the atmosphere may be made up of all the fugitive elements that have been added by inorganic processes over the history of the solar system. Clouds of sulfur compounds including SO_2 and H_2SO_4 are observed in the atmosphere of Venus.

Table 8.5 Composition of the Lower Atmosphere of Mars and Venus

| Gas | Composition (Volume) | |
	Mars*	Venus†
CO_2	95.32%	91.6%
N_2	2.7%	13.5%
Ar	1.6%	20 ppm
O_2	0.13%	60 ppm
CO	0.07%	—
H_2O	0.03% (variable)	0.1%
Ne	2.5 ppm	6–10 ppm
Kr	0.3 ppm	—
Xe	0.08 ppm	—
O_3	0.03 (variable)	—
SO_2	—	180 ppm
pressure	7.6 mbar	91 bar

*From Owen et al., J. Geoph. Res. **82**, 4637 (1977).
†From Oyama et. al. Science 203, 802 (1979).

An interesting comparison of the atmosphere of the earth and Venus may be made by taking all of the carbon and nitrogen of the sedimentary rocks of Earth's crust and adding them to the atmosphere as carbon dioxide and molecular nitrogen. When this is done, we show that the mass of the atmosphere may be increased 6 to 22 times and that, like the atmosphere of Venus, carbon dioxide would now make up over 90% of its mass.

Because of the small amount of nitrogen contained in sedimentary rocks, the absolute abundance of nitrogen only doubles and it becomes a minor constituent compared to the carbon dioxide. This overall similarity between Venus and Earth indicates that they may share a similar chemical history with respect to their fugitive elements. However, on Earth life evolved to produce an oxygen atmosphere. The overwhelming abundance of carbon dioxide in the atmosphere of Venus coupled with the greenhouse effect has caused it to have an extremely high surface temperature, prohibiting the condensation of water and the control of carbon dioxide by the solution of CO_2 in the hydrosphere with the subsequent precipitation of carbonates. Urey has suggested that the type of chemical reaction $MgSiO_3 + CO_2 = MgCO_3 + SiO_2$ requires the presence of liquid water to act as a catalyst. The high surface temperature of Venus precludes the presence of this liquid water and the consequent removal of CO_2 from its atmosphere.

Atmospheric Additions During Geological Time

The following additions have been made to the atmosphere during geological time:

1. Gases released by the crystallization of magmas.
2. Oxygen produced by photochemical dissociation of water vapor.

3.   Oxygen produced by photosynthesis.
4.   Helium from the radioactive decay of uranium and thorium.
5.   Argon from the radioactive decay of potassium.
6.   Gases added by the solar wind, comets, and meteorites.

Carbon dioxide and a very small amount of methane are contributed to the atmosphere by the vital activity and decay of organisms, but this material is cyclic in nature since it was ultimately derived from the atmosphere.

The contribution of igneous activity to the atmosphere is considerable. Water vapor is by far the most abundant volcanic gas, but it is rapidly condensed and added to the hydrosphere along with other readily soluble gases produced by volcanism, such as HCl, HF, H_2S, and SO_2. Next to water, carbon dioxide is probably the most important contribution of igneous activity to the atmosphere. Much of the H_2O and CO_2 is secondary, having been picked up from surrounding rocks during the ascent of the magma, or recycled material carried down in subduction zones. Igneous activity is evidently the source of the CO_2 needed to replace that locked up as carbonate minerals and free carbon during geological time. Rubey (1951) in a classic paper carefully evaluated the contribution of igneous activity to the atmosphere and hydrosphere, and his results are given in Table 8.6. The excess volatiles unaccounted for by rock weathering must be the amounts present in the primeval atmosphere and hydrosphere, plus that added by igneous activity during geological time. From Rubey's data it seems reasonable to assume that the latter quantity far exceeds that originally present.

Some nitrogen has also been added to the atmosphere from magmatic sources during geological time. It is usual to ascribe the nitrogen recorded in volcanic gases to atmospheric contamination, and most of it probably originates thus. However, analyses indicate nitrogen as a constituent of igneous rocks and so presumably of magmatic gases. The nitrogen content of igneous rocks of all compositions, from ultrabasic to acidic, is remarkably constant, averaging about 0.04 cm^3 N_2/g, or 0.005 wt %. The nitrogen is largely present in chemical combination because an apprecia-

Table 8.6   Estimated Quantities (in Units of 10^{20} Grams) of Volatile Materials Now at or Near the Earth's Surface

	H_2O	Total C as CO_2	Cl	N	S	H, B, Br, Ar, F, etc.
In present atmosphere, hydrosphere, and biosphere	14,600	1.5	276	39	13	1.7
Buried in ancient sedimentary rocks	2,100	920	30	4.0	15	15
Total	16,700	921	306	43	28	16.7
Supplied by weathering of crystalline rocks	130	11	5	0.6	6	3.5
Excess volatiles unaccounted for by rock weathering	16,600	910	300	42	22	13

Table 8.7 Production and Use of Oxygen

Total estimated production	
By photosynthesis in excess of decay	181×10^{20} gm
By photodissociation of water vapor followed by hydrogen escape	1
Total	182×10^{20} gm
Total estimated use	
Oxidation of ferrous iron to ferric iron during weathering	14
Oxidation of S^{-2} to SO_4^{-2} during weathering	12
Oxidation of volcanic gases	
CO to CO_2	10
SO_2 to SO_3	11
H_2 to H_2O	<150
Total	$<197 \times 10^{20}$ gm
Free in atmosphere	$\sim 12 \times 10^{20}$ gm

ble amount is driven off as ammonia when the rocks are fused with Na_2CO_3. Chemical combination of the nitrogen is borne out by the nitrogen: argon ratio, which is 2000 in igneous rocks as against 120 in the atmosphere.

A balance sheet for the production and consumption of atmospheric oxygen has been provided by Holland (Table 8.7). The photochemical dissociation of water vapor in the upper atmosphere, with the subsequent escape of hydrogen into outer space, is a process that has been adding to the earth's supply of free oxygen throughout its history. However, the estimated amount produced in this way is insignificant compared with that contributed by photosynthesis.

The total amount of oxygen released by photosynthesis can be calculated if we know the amount of organic carbon that has been "fossilized" in the sedimentary rocks. Rubey estimates that the total organic carbon in the sedimentary rocks is the equivalent of 250×10^{20} g CO_2, which would correspond to the photosynthetic liberation of 181×10^{20} g oxygen.

No reasonable explanation of the organic carbon of sediments other than a biological origin has been proposed, so we may assume 181×10^{20} g oxygen has been liberated by photosynthesis and not reconverted to carbon dioxide by decay or respiration. As indicated earlier the clues for the early buildup of oxygen in the atmosphere come from the interrelations of this atmosphere with sedimentary rocks and the biosphere. The fossil record provides evidence for the period of the rise of photosynthetic organisms. The production of oxygen by the decomposition of water molecules by ultraviolet light is limited to about 10^{-3} of the amount now in the atmosphere. In this reaction the water molecule is split by photodissociation into OH and H radicals. Many radicals recombine to water but some hydrogen is lost by diffusion into space, thereby producing an oxygen buildup. A limiting factor to the buildup of oxygen by this process is the formation of ozone, O_3. Ozone absorbs ultraviolet light at the wavelengths needed for photodissociation and thus prevents the ultraviolet radiation from decomposing the water molecules in the lower atmosphere.

The low oxygen content in the early atmosphere would have allowed the ultraviolet radiation to reach the earth's surface and produce a lethal environment. Ancient

sediments have been found that are less oxidized then their present-day counterparts. In these sediments there is evidence that compounds of uranium, sulfur, manganese, and iron were in a reduced state. All are readily oxidized elements under present atmospheric conditions. Uraninite, UO_2, of apparent detrital origin has been found in sediments from South Africa more than 1.8×10^9 years old. Likewise detrital pyrite has been found in sediments of similar age. Under present atmospheric oxygen levels both would undergo rapid weathering. It appears that around 2×10^9 years ago the earth's atmosphere was not oxidizing.

In view of the lethal radiation at the surface, it has been proposed that photosynthetic organisms began to develop under enough water to provide suitable shielding. These organisms would produce oxygen, much of which would be used up in oxidizing reduced chemical species but some of which would diffuse into the atmosphere. When the oxygen reached 1% of the present atmospheric level, developing organisms could change over from fermentation to respiration. This has been referred to as the "Pasteur point." By this time the ultraviolet light would be well enough shielded so that only a few centimeters of water was needed for protection and life could undergo an explosive evolutionary increase. Cloud and others have suggested that the Precambrian banded iron formations were formed at this stage. These sediments, again about 2.7 to 1.8×10^9 years old, contain both ferric and ferrous iron. It has been proposed that ferrous iron weathered from rocks in an essentially oxygen-free terrestrial environment was oxidized and precipitated when it was washed into the plant-containing oxygenated water. It is also important to note that calcite and dolomite in the early sediments indicate a neutral rather than hydrogen-rich reducing environment, since under the latter conditions CH_4 rather than CO_2 would be the dominant carbon compound.

When the oxygen released by photosynthetic activity reached 10^{-1} of the present atmospheric level about 4×10^8 years ago (Silurian) the earth's surface was shielded from harmful ultraviolet radiation. Life could now progress to the water's surface and to land and the oxygen content could grow until it reached the present level.

Radioactive daughter products are also produced in the earth and degassed into its atmosphere. The rate of production of helium is 1.16×10^{-7} cm^3/g of uranium and 2.43×10^{-8} cm^3 per gram of thorium per year. A simple calculation based on the amounts of uranium and thorium in the earth and an age of 4.5×10^9 years shows that the quantity of helium liberated by radioactive decay during geological time is greater than that now present in the atmosphere. This difference is explained by the loss of helium from the earth into interplanetary space, which is discussed in the next section.

Argon has three isotopes, ^{36}Ar, ^{38}Ar, and ^{40}Ar, of which ^{40}Ar is by far the most abundant (the percentages of ^{36}Ar, ^{38}Ar, and ^{40}Ar in air are 0.307, 0.061, and 99.632). The amount of argon on the earth is anomalously high when compared with that of the other inert gases. This is evidently due to the production of ^{40}Ar by the radioactive decay of ^{40}K throughout geologic time. Much of the argon thereby produced has probably remained occluded in the potassium minerals where it originated, but sufficient has been liberated to the atmosphere to give it a comparatively high argon content.

Atmospheric Losses During Geological Time

Atmospheric losses during geologic time may be physical, as in the loss of hydrogen or helium to space by diffusion, or chemical, by reaction between the atmosphere and the hydrosphere, biosphere, or lithosphere. Some important loss mechanisms can be summarized under the following categories:

1. Loss of oxygen by oxidation of hydrogen to water, ferrous to ferric iron, sulfur compounds to sulfates, manganese compounds to manganese dioxide, and similar reactions.
2. Loss of carbon dioxide by the formation of coal, petroleum, and disseminated carbon from the death and burial of organisms.
3. Loss of carbon dioxide by the formation of calcium and magnesium carbonates.
4. Loss of nitrogen by the formation of oxides of nitrogen in the air and by the action of nitrifying bacteria in the soil.
5. Loss of hydrogen and helium by escape from the earth.

Removal of oxygen is observable wherever weathering is taking place. Ferrous iron oxidizes readily to ferric iron under most surface conditions, as is evidenced by the yellow or red colors produced by ferric oxides in many sediments. Figures for the average composition of igneous rocks show more FeO than Fe_2O_3, whereas for sedimentary rocks the relationship is reversed. Iron is the principal consumer of oxygen during weathering, but manganese is converted from the bivalent to the quadrivalent state, and oxidation converts sulfides into sulfates or free sulfuric acid. From Table 8.7, however, it is clear that the principal consumption of free oxygen has been through the oxidation of volcanic gases, especially hydrogen.

The loss of carbon dioxide from the atmosphere by deposition as carbonate and organic carbon in sedimentary rocks was estimated by Rubey as totaling 920×10^{20} g. Wickman (1956) has published some revised figures. He places the amount of carbonate carbon per square centimeter of the earth's surface at 2420 ± 560 g and of organic carbon at 700 ± 200 g. Taking the figure of 3100 g/cm^2 for the total amount of carbon transferred from the atmosphere to the sedimentary rocks, this is equal to a total of 158×10^{20} g carbon, or 580×10^{20} g CO_2. The latter figure is of the same order of magnitude as Rubey's estimate but considerably lower. The figures show clearly that the amount of carbon dioxide deposited in sedimentary rocks far exceeds the amount in the present atmosphere, hydrosphere, and biosphere (about 1.5×10^{20} g), and thus indicate that large amounts of carbon dioxide must have been released or recycled from magmatic sources throughout geological time to maintain organic activity. Wickman's figures show, in addition, that far more carbon dioxide has been removed as limestone and dolomite than as coal or other organic carbon.

Nitrogen is removed from the air by both organic and inorganic processes. The organic processes include nitrogen-fixing microorganisms living in the root nodules of certain plants, soil organisms (both aerobes and anaerobes), and some blue-green algae. Inorganic processes produce oxides of nitrogen by electrical discharges and photochemical reactions in the atmosphere. The amount of organically fixed nitrogen

greatly exceeds that fixed by inorganic processes; Hutchinson (1954) carefully ana-
lyzed the geochemical cycle of nitrogen and estimated biological fixation at 0.008 to
0.07 mg N_2/cm^2 per year for the land surfaces of the earth, nonbiological fixation at
not more than 0.0035 mg N_2/cm^2 per year. Much of this nitrogen is eventually
returned to the atmosphere by the decay of organic matter. Some, however, remains
in the sediments and in a few places may be sufficiently concentrated to form nitro-
genous deposits, such as the Chilean nitrates; lower concentrations of nitrogen are
found in guano deposits. Hutchinson gives a figure of 67 to 108 g/cm^2 of the earth's
surface for the amount of fossil nitrogen in sediments; of this fossil nitrogen he esti-
mates that about 8 g/cm^2 have been derived from the weathering of igneous rocks,
so that 60 to 100 g/cm^2 have been abstracted from the atmosphere. Assuming a
figure of 80 g/cm^2, this amounts to a total of 4.8×10^{20} g nitrogen removed from
the atmosphere during geological time, a very small amount compared to that of
carbon dioxide.

The loss of gases by escape from the earth or any other planet depends on the
strength of its gravitational field and the mean square velocity of the gas molecules.
The mean square velocity is the velocity whose square is equal to the mean of the
squares of the individual velocities; it varies inversely as the square root of the molec-
ular weight of the gas and directly as the square root of the absolute temperature.
At 0° the mean square velocity is 1.84 km/sec for hydrogen, 1.31 for helium, 0.62
for water vapor, 0.49 for nitrogen, 0.46 for oxygen, and 0.39 for carbon dioxide. At
100° these velocities are increased by 17%. For any body there is a so-called velocity
of escape which may be calculated from the formula

$$V^2 = \frac{2GM}{R}$$

where V is the velocity, G is the gravitational constant, M is the mass of the body,
and R its radius. For the earth the gravitational constant is 6.67×10^{-8}, the mass
5.97×10^{27} g, and its radius 6.37×10^8 cm; therefore, the velocity of escape is 11.3
km/sec. Even if the mean velocity of the molecules is considerably less than the
velocity of escape, an atmosphere will gradually be lost by the escape of fast-moving
molecules from its extreme upper region, where the free paths of the molecules are
so long that they stand a chance of getting away without being stopped by collisions.
Jeans showed that an atmosphere will be stable throughout geological time ($> 10^9$
years) if the mean square velocity of the molecules is less than one-fifth the velocity
of escape. On this basis no loss of gases into interplanetary space should have taken
place since the earth cooled to its present temperature, because even for hydrogen
the mean square velocity is considerably less than the velocity of escape. The situa-
tion is, however, not quite as simple as it may appear. First, hydrogen and helium
atoms colliding with metastable oxygen atoms in the upper atmosphere acquire suf-
ficient momentum for their velocities to exceed the escape velocity. In addition, tem-
peratures in the upper atmosphere are considerably higher than at the earth's surface
(recent estimates suggest temperatures between 200 and 1200° at altitudes of 120
to 700 km, as a result of absorption of ultraviolet radiation); this sufficiently increases
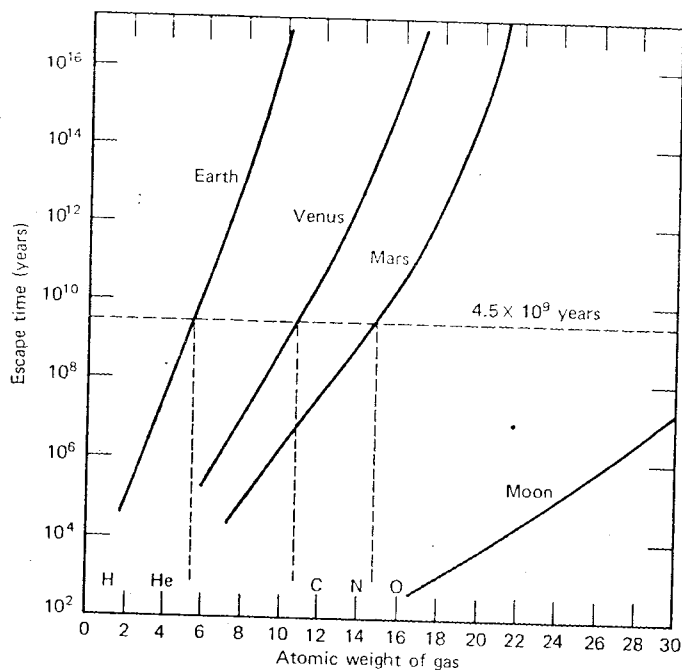
Figure 8.5 Calculated effective time of escape of gases for earth,
Venus, Mars and the moon. (After Jastrow and
Rasool in *Introduction to Space Science,* W. N. Hess
ed; Gordon and Breach, New York 1965, p 676.)

the mean square velocities of hydrogen and helium to permit their escape. The fact
that no more than one-tenth the amount of helium supplied by radioactive disinte-
gration during geological time is present in the atmosphere shows that such escape
has occurred. Venus, which is closer to the sun and of a higher temperature than
Earth, and Mars, although cooler but of a smaller size, both will lose gases more
readily than Earth. The calculated effective time of escape of gases for these three
planets is shown in Figure 8.5.

CONSTANCY OF ATMOSPHERIC COMPOSITION

Geological evidence indicates a uniformity in climatic and biological conditions since
early Paleozoic times that could not have existed had the atmosphere been subject
to marked changes in composition. Nevertheless, minor changes in atmospheric com-
position have been suggested as possibly responsible for climatic variations. For
example, the onset of glacial periods has been ascribed to a decrease in carbon diox-
ide, since this gas is a selective absorbent for solar radiation. On the other hand,
water vapor performs a similar function, and its presence in the atmosphere in large
amounts far outweighs the climatic significance of CO_2. The present opinion is that

only a small part of climatic variations can be attributed to varying CO_2 content of the atmosphere.

Eskola made the interesting suggestion of a correlation between periods of rapid organic evolution and periods during which carbon dioxide was available in greatest amounts. He pointed out that rapid organic evolution has coincided with worldwide orogeny. During orogenic periods igneous activity is at a maximum; hence the addition of carbon dioxide to the atmosphere also reaches a maximum. The increased supply of carbon dioxide would stimulate organic activity and be at least a contributory factor in the evolution of higher forms of life.

Considerable interest has been aroused by the prospect of a marked increase of carbon dioxide in the atmosphere as a result of the greater use of fossil fuels in recent years. The latest figure for the amount of coal mined annually is 1.63×10^9 tons; assuming an average of 78% C, this would produce 4.67×10^{15}g CO_2. For oil the annual production figure is 3.4×10^9 barrels (1 barrel = 160 liters); assuming a density of 0.9 and an average of 85% C, this would produce 1.53×10^{15}g CO_2. The total CO_2 produced by the burning of the annual production of coal and oil is 6.2 $\times 10^{15}$ g, or about 1/300 of the amount in the atmosphere today. This might suggest that at the present rate of consumption of fossil fuels atmospheric carbon dioxide will be doubled in 300 years. However, in this connection the importance of the hydrosphere as a reservoir of carbon dioxide should be emphasized; its significance has been discussed by Broecker and co-workers (1979). Sea water contains 20 g CO_2/cm^2 of the earth's surface, as against 0.4 g/cm^2 in the atmosphere. Oceanic and atmospheric carbon dioxide are interdependent, the former being a function of the partial pressure of CO_2 in the atmosphere. Thus to double the partial pressure of carbon dioxide in the atmosphere would require the addition of much more than is now present therein, because most of that added would be absorbed by the ocean; similarly, to decrease the carbon dioxide in the atmosphere by one-half would require the removal of many times the present content. It is apparent that the oceans, by controlling the amount of atmospheric CO_2, play a vital part in maintaining stable conditions suitable for organic life on the earth.

SELECTED REFERENCES

Abelson, P. H. (1966). Chemical events on the primitive earth. *Proc. Nat. Acad. Sci.* **55**, 1365–1372. A review of the nature of the primitive atmosphere and ocean in the light of geological and geophysical information.

Berner, R. A. (1971). Worldwide sulfur pollution of rivers. *J. Geophs. Res.* **76**, 6597–6600. An estimate is made that about 28% of sulfur in rivers is contributed by pollution.

Brancazio, P. J. and A. G. W. Cameron (eds.) (1964). *The origin and evolution of atmospheres and oceans.* 314 pp. John Wiley & Sons, New York. The proceedings of a conference on the processes responsible for the production and subsequent development of planetary atmospheres and oceans; also includes a reprint of Rubey's article "Geologic history of sea water" (1951).

Broecker, W. S., T. Takahashi, H. J. Simpson, and T.-H. Peng (1979). Fate of fossil fuel carbon dioxide and the global carbon budget. *Science* **206**, 409–418. A quantitative discussion of the carbon dioxide cycle.

Chamberlin, J. W. (1978). *Theory of planetary atmospheres*. 330 pp. Academic Press, New York. Considers both chemistry and physics of atmospheres.

Cloud, P. (1976). Beginnings of biospheric evolution and their biogeochemical consequences. *Paleobiology* **2**, 351–387.

Heicklen, J. (1976). *Atmospheric chemistry*. 406 pp. Academic Press, New York.

Holland, H. D. (1962). Model for the evolution of the earth's atmosphere. *Geol. Soc. Am.*, Buddington Vol., 447–477. A carefully reasoned account of the processes by which the present atmosphere was developed from a primordial reducing atmosphere.

——— (1978). The chemistry of the atmosphere and oceans. 351 pp. John Wiley & Sons, New York. A complete review of the volatiles at the earth's surface.

Holser, W. T. and I. R. Kaplan (1966). Isotope geochemistry of sedimentary sulfates. *Chem. Geol.* **1**, 93–135.

Hutchinson, G. E. (1954). The biogeochemistry of the terrestrial atmosphere. In *The earth as a planet*, G. P. Kuiper, (ed.). Chapter 8 (pp. 371–433). University of Chicago Press, Chicago. An exhaustive review of the factors that influence the composition of the atmosphere, with evaluations of their quantitative effects.

Junge, C. E., M. Schidlowski, R. Eichmann, and H. Pietrek (1975). Model calculations for the terrestrial carbon cycle: Carbon isotope geochemistry and evolution of photosynthetic oxygen. *J. Geophys. Res.* **80**, 4542–4552. A mathematical evaluation of this important question.

Junge, C. E. and R. T. Werby (1958). The concentration of chloride, sodium, potassium, calcium and sulfate in rain water over the United States. *J. Meteorol.* **15**, 417–425.

McEwan, M. J. and L. F. Phillips (1975). *Chemistry of the atmosphere*. 301 pp. Halsted Press, John Wiley & Sons, New York.

Maxwell, J. R., C. T. Pillinger, and G. Eglinton (1971). Organic geochemistry. *Q. Rev. (London)* **25**, 571–628.

Meadows, A. J. (1972). The atmospheres of the earth and the terrestrial planets. *Physics Reports* (Sec. C. of Physics Letters) **5**, 197–236.

Rubey, W. W. (1951). Geologic history of sea water. *Bull. Geol. Soc. Amer.* **62**, 1111–1147. This paper contains a great amount of data on the gases contributed to the atmosphere and the hydrosphere by magmatic activity and discusses the geological implications.

——— (1955). Development of the hydrosphere and atmosphere, with special reference to probable composition of the early atmosphere. *Geol. Soc. Amer. Spec. Paper* **62**, 631–650. The different hypotheses for the origin of the earth's atmosphere and hydrosphere are considered in the light of physicochemical principles and geological evidence.

Scientific American (1970). *The biosphere*, 134 pp. A series of articles, many on the geochemical cycles of atmospheric constituents.

Urey, H. C. (1952). *The planets*. 245 pp. Yale University Press, New Haven. One topic discussed in this book is the origin and composition of the earth's primitive atmosphere.

Wickman, F. E. (1956). The cycle of carbon and the stable carbon isotopes. *Geochim. Cosmochim. Acta* **9**, 136–153. The amounts of carbonate carbon and organic carbon in sedimentary rocks is estimated from measurements of $^{12}C/^{13}C$ fractionation.

9 THE NATURE OF THE HYDROSPHERE

The hydrosphere is the discontinuous shell of water—fresh, salt, and solid—at the surface of the earth. It comprises the oceans with their connected seas and gulfs, the lakes, the waters of the rivers and streams, ground water, and snow and ice. The oceans are clearly of first magnitude, for they cover an area of 361×10^6 km^2, or 70.8% of the earth's surface. From the mean depth of 3800 m, the volume of the ocean waters can be calculated to be 1370×10^6 km^3. At the surface the density of sea water of normal salinity at 0°C is 1.028, and it increases with depth because of the slight compressibility of water under increased pressure. It is more difficult to arrive at an accurate measure of water in the other parts of the hydrosphere but an estimate is given in Table 9.1. Seawater thus comprises about 97% of the weight of the hydrosphere and its composition can therefore be taken without serious error as giving an average composition of the hydrosphere.

The circulation of the waters in the hydrosphere provides a mechanism for geo-chemical changes through weathering and transport of dissolved and solid materials.

Table 9.1 Distribution of Water on the Earth

	Volume, km^3	Mass, g
Seawater	1.37×10^9	1.41×10^{24}
Fresh water lakes	1.3×10^5	1.3×10^{20}
Saline lakes and inland seas	1.0×10^5	1.0×10^{20}
Rivers	1.3×10^3	1.3×10^{18}
Groundwater	8.4×10^6	8.4×10^{21}
Icecaps	2.9×10^7	2.9×10^{22}
Water vapor	1.3×10^4	1.3×10^{19}

The annual evaporation of water from the ocean of 3.5×10^5 km³ and land area of 7×10^4 km³ is returned as precipitation of 3.2×10^5 km³ to the oceans and 1.0×10^5 km³ to the land areas. About 68% of the land area is drained to the Atlantic Ocean but the constancy of ocean composition indicates that seawater mixing erases all but local variation in its chemical composition.

THE COMPOSITION OF SEAWATER

Two arbitrarily defined quantities, the chlorinity and the salinity, are used by ocean-ographers in discussing the composition of seawater. The chlorinity is determined by the precipitation of the halides with a silver salt and is essentially the total amount in grams of chloride, bromide, and iodide contained in 1 kg of seawater, assuming that the bromide and iodide have been replaced by chloride. The salinity is a specif-ically defined quantity, slightly less than the total weight of dissolved solids per kilo-gram, and can be calculated from the chlorinity or determined from a measurement of physical properties such as density, refractive index, or electrical conductivity. Both chlorinity and salinity are customarily expressed in grams per kilogram (g/kg) of seawater or parts per thousand (‰). In the open ocean the salinity averages about 35‰ with a regional range between 32 and 37‰, but rises to as much as 41‰ in the Red Sea and the Persian Gulf, where evaporation is high and inflow and precipitation low, and decreases where fresh water from rivers or melting ice enters the ocean. However, in all samples of seawater the relative proportions of the major ions are practically constant, thus the determination of one constituent provides a measure of the others. Because of the homogeneity of seawater among the most accurate data of geochemistry are those regarding the ocean.

Early knowledge of the average composition of seawater was provided by Dittmar, who in 1884 carefully analyzed 77 water samples, representative of all oceans and taken both from the surface and from the depths. These samples had been collected on the voyage around the world of *H.M.S. Challenger* (1872–1876). He determined the halides, sulfate, magnesium, calcium, and potassium. On composite samples he found the ratio of bromine to chlorine and estimated the carbonate. Sodium was calculated by difference by subtracting the sums of the chemical equivalents of the negative and positive ions. This procedure was followed because Dittmar was unable to achieve satisfactory direct determinations for sodium. Dittmar's work showed that there are no significant regional differences in the relative composition of seawater, and his average values can be used to represent the ratios between the major dis-solved constituents. Since 1884 much research has been devoted to the chemical com-position of seawater. Nevertheless, the figures obtained by Dittmar agree closely with those accepted today as the best available, a remarkable tribute to his work.

The figures for the major constituents are given in Table 9.2 and are referred to a chlorinity of 19‰, which is taken as the standard concentration of seawater. Because of the constancy of the relative proportions of the major constituents any one may be used as a measure of the others, and chlorinity is the most readily deter-mined. The complexity of seawater makes it impossible by direct chemical analysis to determine the total quantity of dissolved solids in a given sample. Furthermore,

Table 9.2   The Major Dissolved
Constituents of Seawater

Ion	Cl = 19‰	Percent
Cl	18.980	55.05
Br	0.065	0.19
SO$_4$	2.649	7.68
HCO$_3$	0.140	0.41
F	0.001	0.00
H$_3$BO$_3$	0.026	0.07
Mg	1.272	3.69
Ca	0.400	1.16
Sr	0.008	0.03
K	0.380	1.10
Na	10.556	30.61
Total	34.477	99.99

reproducible results cannot be obtained by evaporating seawater and weighing the residue, because some of the components, particularly chloride, are lost in the the last stages of drying. Hence, the use of indirect methods based on the chlorinity factor.

The elements given in Table 9.2 are the major dissolved materials and show little relative variation in the open sea. Minor and trace elements present may be separated into two groups: those involved in biological cycles and which show variations with respect to depth and ocean basin or both, and those whose reported variations in concentration are independent of salinity and not obviously depth or basin dependent. Table 9.3 gives available data on the amounts of all elements in seawater. For C, N, and O separate concentrations are given for individual chemical species because of important differences in their distribution. The elements in category A are those directly proportional to salinity; those in category B are the nutrient and other elements involved in biochemical cycles; and those in category C are those not involved in biochemical cycles and appear to be independent of the salinity.

It should be realized that the quality of the chemical data available is uneven, because some figures are based on the analysis of a single sample and many of the samples were taken from inshore surface waters, which for minor and trace elements may not be representative of the ocean as a whole. The concentrations of elements in group B fluctuate from place to place as a result of biological activity. These fluctuations are often related to the depth at which the sample was taken, biological activity being greatest in surface and near-surface waters. Calcium may be relatively diminished in surface layers through abstraction by organisms. Silica is also removed from surface waters in this way, and its content in seawater generally shows a regular increase with depth. The distribution of phosphorus is greatly affected by organisms; this element is markedly enriched in the deeper parts of the ocean as a result of the dissolution of dead organic matter. A typical vertical phosphorus profile is shown in Figure 9.1.

Table 9.3 Elements Present in Seawater and Their Geochemical Characteristics

Element	Seawater Concentration (μg/liter)	Principal Dissolved Species	Category	Concentration Dissolved in Stream Waters (μg/liter)	Residence Time in Ocean (yr)
H	1.1×10^8	H_2O	A	—	—
He	7×10^{-3}	He(gas)	A	—	—
Li	1.7×10^2	Li^+	A	3	2.3×10^6
Be	6×10^{-4}	—	—	—	—
B	4.5×10^3	$B(OH)_3$, $B(OH)_4^-$	A	10	1.8×10^7
C	2.8×10^4	HCO_3^-, CO_3^{-2}	A	—	—
C(org.)	1×10^2	—	A	—	—
N	1.5×10^4	N_2 (gas)	A	—	—
N	6.7×10^2	NO_3^-	B	—	—
O	8.8×10^8	H_2O	A	—	—
O	6×10^3	O_2	B	—	—
O	1.8×10^6	SO_4^{-2}	A	—	—
F	1.3×10^3	F^-	A	100	5.2×10^5
Ne	0.12	Ne(gas)	A	—	—
Na	1.1×10^7	Na^+	A	6,300	6.8×10^7
Mg	1.3×10^6	Mg^{+2}	A	4,100	1.2×10^7
Al	1	—	—	400	1.0×10^2
Si	3×10^3	$Si(OH)_4$, $SiO(OH)_3^-$	B	6,500	1.8×10^4
P	90	HPO_4^{-2}, $H_2PO_4^-$, PO_4^{-3}	B	20	1.8×10^5
S	9.0×10^5	SO_4^{-2}	A	—	—
Cl	1.9×10^7	Cl^-	A	1,800	1×10^8
Ar	4.5×10^2	Ar(gas)	A	—	—
K	3.9×10^5	K^+	A	2,300	7×10^6
Ca	4.1×10^5	Ca^{+2}	A	15,000	1.0×10^6
Sc	$<4 \times 10^{-3}$	$Sc(OH)_3^0$	—	0.004	$<4 \times 10^4$
Ti	1	$Ti(OH)_4^0$	—	3	1.3×10^4
V	2	$VO_2(OH)_3^{-2}$	—	0.9	8.0×10^4
Cr	0.5	CrO_4^{-2}, Cr^{+3}	—	1	2.0×10^4
Mn	2	Mn^{+2}	C	7	1.0×10^4
Fe	3	—	—	670	2.0×10^2
Co	0.4	Co^{+2}	C	0.1	1.6×10^5
Ni	7	Ni^{+2}	C	0.3	9.0×10^4
Cu	3	Cu^{+2}	C	7	2×10^4
Zn	10	Zn^{+2}	C	20	2×10^4
Ga	3×10^{-2}	—	—	0.09	1×10^4
Ge	7×10^{-2}	$Ge(OH)_4$	—	—	—
As	2.6	$HAsO_4^{-2}$, $H_2AsO_4^-$	—	2	5×10^4
Se	9×10^{-2}	SeO_4^{-2}	C	0.2	2×10^4
Br	6.7×10^4	Br^-	A	20	1×10^8
Kr	0.2	Kr(gas)	A	—	—
Rb	1.2×10^2	Rb^+	A	1	5×10^6
Sr	8×10^3	Sr^{+2}	A	70	4×10^6
Y	1×10^{-3}	$Y(OH)_3^0$	C	—	—
Zr	3×10^{-2}	—	—	—	—
Nb	0.01	—	—	—	—
Mo	10	MoO_4^{-2}	A	0.6	7×10^5
Ru	—	—	—	—	—
Rh	—	—	—	—	—
Pd	—	—	—	—	—
Ag	0.3	$AgCl_2^-$	C	0.3	4×10^4

Table 9.3 Elements Present in Seawater and Their Geochemical Characteristics (*continued*)

Element	Seawater Concentration (μg/liter)	Principal Dissolved Species	Category	Concentration Dissolved in Stream Waters (μg/liter)	Residence Time in Ocean (yr)
Cd	0.1	Cd^{+2}	—	—	—
In	<20	—	—	—	—
Sn	0.8	—	—	—	—
Sb	0.3	—	C	2	7,000
Te	—	—	—	—	—
I	60	IO_3^-, I	A	7	4×10^5
Xe	5×10^{-2}	Xe(gas)	A	—	—
Cs	0.3	Cs^+	A	0.02	6×10^5
Ba	20	Ba^{+2}	C	20	4×10^4
La	3×10^{-3}	$La(OH)_3^0$	C	0.2	6×10^2
Ce	1×10^{-3}	$Ce(OH)_3^0$	C	—	—
Pr	0.6×10^{-3}	$Pr(OH)_3^0$	C	—	—
Nd	3×10^{-3}	$Nd(OH)_3^0$	C	—	—
Sm	0.5×10^{-3}	$Sm(OH)_3^0$	C	—	—
Eu	0.1×10^{-3}	$Eu(OH)_3^0$	C	—	—
Gd	0.7×10^{-3}	$Gd(OH)_3^0$	C	—	—
Tb	1.4×10^{-3}	$Tb(OH)_3^0$	C	—	—
Dy	0.9×10^{-3}	$Dy^2(OH)_3^0$	C	—	—
Ho	0.2×10^{-3}	$Ho(OH)_3^0$	C	—	—
Er	0.9×10^{-3}	$Er(OH)_3^0$	C	—	—
Tm	0.2×10^{-3}	$Tm(OH)_3^0$	C	—	—
Yb	0.8×10^{-3}	$Yb(OH)_3^0$	C	—	—
Lu	0.1×10^{-3}	$Lu(OH)_3^0$	C	—	—
Hf	$<8 \times 10^{-3}$	—	—	—	—
Ta	$<3 \times 10^{-3}$	—	—	—	—
W	0.1	WO_4^{-2}	—	0.03	1.2×10^5
Re	0.008	—	—	—	—
Os	—	—	—	—	—
Ir	—	—	—	—	—
Pt	—	—	—	—	—
Au	1×10^{-2}	$AuCl_2^-$	C	0.002	2×10^5
Hg	0.2	$HgCl_4^{-2}$, $HgCl_2^0$	C	0.07	8×10^4
Tl	<0.1	Tl^+	—	—	—
Pb	0.03	$PbCl_3^-$, $PbCl^+$, Pb^{+2}	C	3	4×10^2
Bi	0.02	—	—	—	—
Po	—	—	—	—	—
At	—	—	—	—	—
Rn	6×10^{-13}	Rn(gas)	—	—	—
Ra	1×10^{-7}	Ra^{+2}	C	—	—
Ac	—	—	—	—	—
Th	$<5 \times 10^{-4}$	$Th(OH)_4^0$	—	0.1	<200
Pa	2.0×10^{-6}	—	—	—	—
U	3	$UO_2(CO_3)_3^{-4}$	A	0.04	3×10^6
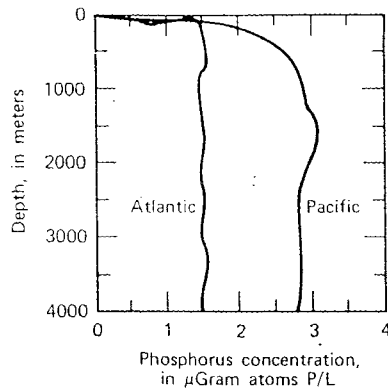
Figure 9.1 Vertical distribution of typical inorganic phosphate concentration in the North Atlantic and Pacific oceans. Other nutrient components such as nitrate show similar curves with depletion near the surface.

Seawater also contains gases in solution. Since the atmosphere and the ocean are in contact, a relation must exist between the amount of the gases in solution and their partial pressures in the atmosphere. The surface waters are in equilibrium or near equilibrium with the oxygen and nitrogen of the air. It is generally assumed that nitrogen dissolved in seawater does not enter into chemical reactions; hence, its concentration is not subject to appreciable variation. Oxygen on the other hand plays an active part in metabolism and in the decay of organic matter, and its percentage varies considerably from place to place. The atmosphere also regulates the carbon dioxide content of surface waters, but the relationship is complex because carbon dioxide is present in seawater in four distinct chemical species; free carbon dioxide, carbonate ions, bicarbonate ions, and undissociated H_2CO_3. The carbon dioxide content of seawater, which is the most important factor controlling the solubility of $CaCO_3$, is also dependent on the nature and amount of biological activity. Ammonia, argon, helium, and neon have been recorded in seawater. Hydrogen sulfide is often locally present (probably in part as sulfide ions rather than free gas), and may be widespread in stagnant bottom waters. Table 9.4 gives some data on dissolved gases in seawater.

Table 9.4 Dissolved Gases in Seawater*

	Concentration (ml/l)
Oxygen	0–9
Nitrogen	8.4–14.5
Total carbon dioxide	34–56
Argon (residue after removal of N)	0.2–0.4
Helium and neon	1.7×10^{-4}
Hydrogen sulfide	0–22 or more

*Seawater of 19‰ chlorinity at 0° in equilibrium with normal dry atmosphere will contain 8.08 ml/l oxygen and 14.40 ml/l nitrogen.

THE COMPOSITION OF TERRESTRIAL WATERS

Although the total amount of terrestrial waters is insignificant in comparison with the total mass of the hydrosphere, these waters are important geochemically, because they are responsible for most of the weathering and erosion of the land masses and because of their importance to human activities. A knowledge of their amount and composition is clearly essential to an understanding of the evolution of the ocean and as background information for the analysis of human interaction with and pollution of fresh-water systems. The ultimate source of most terrestrial waters is rain, although some magmatic water may be added from volcanism and through thermal springs. Part of the total rainfall runs directly off into streams, part is taken up as ground water and may ultimately reappear as springs, part is retained by the formation of hydrated compounds, and part is returned to the atmosphere. Because of the solvent power of water the runoff from the land is never pure H_2O but always contains dissolved material. The amount of this dissolved material, however, differs greatly from time to time and from place to place. Figure 9.2 illustrates the distribution of various constituents in terrestrial water in the form of cumulative curves. Note that the major dissolved constituent is bicarbonate HCO_3^-, derived from the solution of atmospheric CO_2 in rain water and the addition of CO_2 by organic activity in soils. Other common constituents include Na^+, Ca^{++}, SO_4^{2-}, and Cl^-, in part added by cycled sea salts via precipitation and dry fallout and in part by the weath-
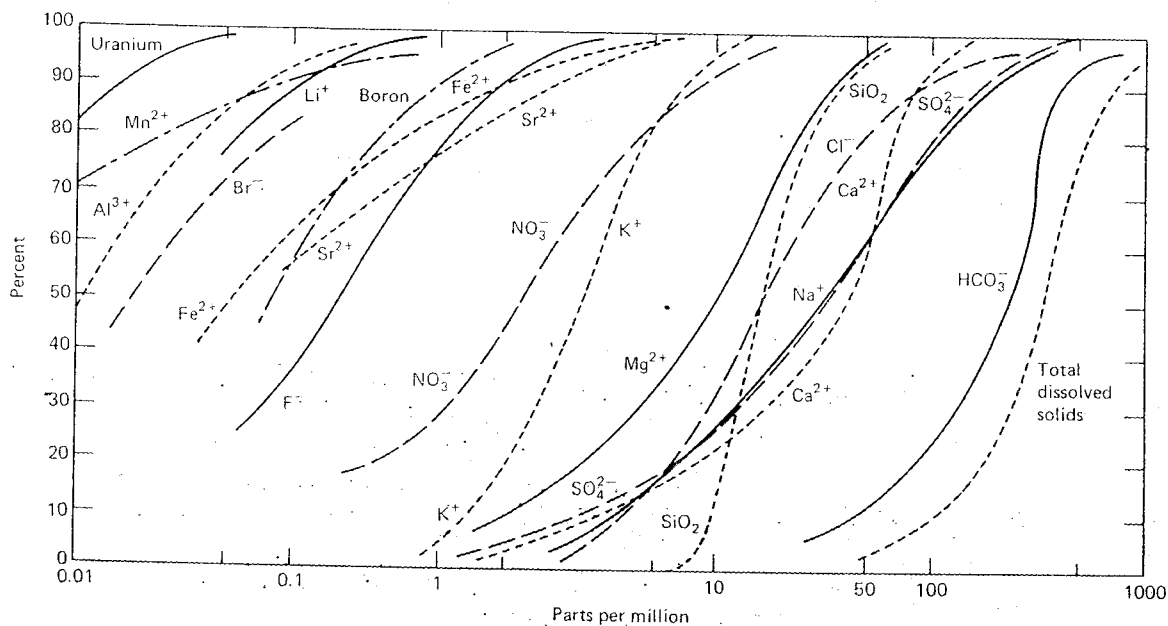
Figure 9.2   Cumulative curves showing the distribution of various constituents in terrestrial water. Data are mostly for the United States from various sources. (S. N. Davies and R. C. N. DeWiest, *Hydrogeology*, John Wiley, 1966.)

ering of rocks. The composition of terrestrial waters in contact with igneous rocks may to the first approximation be considered as due to the reaction of CO_2-rich waters with plagioclase feldspar and dark minerals such as hornblende, pyroxene, and biotite. Typical weathering reactions such as $CaAl_2Si_2O_8 + 2CO_2 + 3H_2O = Al_2Si_2O_5(OH)_4 + Ca^{2+} + 2HCO_3^-$ show the effect of the carbonated water on igneous minerals to produce clay minerals and soluble ions. Also critical to the rate of the dissolution process is the presence of organic compounds to assist the breakdown at mineral surfaces.

It has been estimated that the total annual rainfall on the land areas of the earth amounts to 123×10^{18} g, of which 27.4×10^{18} g drains off to the sea. Clarke estimated that this water carries 27.35×10^{14} g of dissolved material, which gives an average salinity for river water of about 100 ppm. Actual salinities are very variable; analyses of river water show salinities from 13 to 9185 ppm, but figures greater than 1000 ppm are uncommon. Clarke's data show that the concentrations of carbonate, calcium, and magnesium rise rapidly with salinity until limiting values are reached at about 200 ppm; higher salinities are largely due to increases in sodium, sulfate, and chloride. Waters of salinity up to 50 ppm drain areas consisting mainly of igneous or metamorphic rocks, whereas for salinities of 50–200 ppm the drainage is largely from sedimentary rocks; higher figures indicate large-scale human contamination or drainage of arid regions where saline soils are common.

Clarke computed an average composition of the dissolved matter in river water by weighting the composition of the water from different river systems in proportion to the total supply of dissolved material. Revised figures based on more recent data have been provided by Livingstone (1963), and are given in Table 9.5. Livingstone's data indicate a mean salinity of about 130 ppm.

The compositions of individual rivers and lakes can vary appreciably from the average. Table 9.5 indicates the composition of dissolved solids in selected rivers, illustrating the major chemical types Ca-HCO_3, Ca-SO_4, and Na-Cl. Note that the world, discharge-weighted, mean chemical composition lies close to the Ca-HCO_3 type rivers, indicating that the other types make up only a minor portion of the total

Table 9.5 Dissolved Solids: Composition of Various Rivers and Other Waters (ppm)

	HCO_3	SO_4	Cl	Na	Mg	Ca	Fe	SiO_2	Total Dissolved Solids
World average	58.4	11.2	7.8	6.3	4.1	15.0	0.67	13.1	114
Dilute Ca-HCO_3-Amazon River	17	3.0	1.7	1.8	1.1	4.3	0.05	7.0	36
Ca-HCO_3-Mississippi River	101	41	15	11	7.6	34	0.02	5.9	221
Ca-SO_4-Colorado River	183	289	113	124	30	94	0.01	14	853
Na-Cl-Jordan River	238	174	473	253	71	80.	—	—	1310
Seawater	140	2700	19000	11000	1300	410	0.003	3	34,481
Great Salt Lake	—	6680	55480	33170	2760	160	—	—	203,490

Table 9.6   Chemical Composition of Nile
Mud and Mississippi Silt

	Nile	Mississippi
SiO_2	45.10	69.96
Al_2O_3	15.95	10.52
Fe_2O_3	13.25	3.47
MgO	2.64	1.41
CaO	4.85	2.17
K_2O	1.95	2.30
Na_2O	0.85	1.51
CO_2	?	1.40
H_2O	6.70	3.78

Data from Clarke (1924).

discharge. The chemical composition of a river is a function of precipitation amount and composition together with the rock types, relief, and vegetation of the area which it drains. The precipitation that supplies a river basin reaches the river system in two ways: first, by direct runoff and, second, as groundwater through soils and rocks. Groundwater provides a more constant source with respect to amount and composition than does runoff, which may account for seasonal dilution effects in some river systems. Groundwater passing through limestone basins will be high in Ca or if in dolomite high in both Ca and Mg. High percentages of evaporites commonly found in arid regions will contribute high Cl^- and SO_4^{2-} to rivers. The pH of rivers varies from lows of 4 in tropical streams having high organic acid and high dissolved silica contents to values above 9. The average pH value, however, is between 6 and 8.

In addition to dissolved loads rivers may also carry a high suspended load, usually larger than the dissolved load by a factor of about 4 : 1. This suspended material carried to the shores of the ocean is estimated to be from 100 to 300×10^{14} g each year. The chemical composition of this suspended load is difficult to estimate. Table 9.6 gives data on mud and silt from the Nile and Mississippi Rivers compiled by Clarke.

Lake waters generally are similar to those of the rivers that feed them. Exceptions occur in closed basins where evaporation and salt accumulation is high (Table 9.5), in closed lakes closely related to volcanic phenomena, and in lakes of high organic activity often caused by the polluting activities of man. Many closed basins have accumulated salts with compositions similar to seawater, indicating that they were leached from marine evaporates or marine sediments. Volcanic or outgassing activity can produce exotic concentrations in closed systems due to the addition of HF, HCl, or H_2SO_4. In both lakes and the ocean, living organisms depend on the availability of the dissolved nutrients nitrogen and phosphorus. If excess nutrients are added to a lake, rapid organic growth may take place. The natural metabolism of this excess organic matter may use up the dissolved oxygen in the waters, producing an anoxic condition that is hostile to higher forms of life. This process, called eutrophication, may greatly upset the chemical balance of a lake.

GAINS AND LOSSES OF THE OCEAN

Comparison shows that river water and seawater are opposites in chemical character. In seawater Na > Mg > Ca and Cl > SO$_4$ > CO$_3$; in average river water Ca > Na > Mg and CO$_3$ > SO$_4$ > Cl. In addition, over 90% of the chloride and a good deal of the sulfate in river waters is cyclic, being derived ultimately from the oceans via the atmosphere. Hence the average composition must be adjusted for these cyclic salts in discussing its effect on oceanic composition. This correction has been made in Table 9.7 by assuming that all the chloride in river water is cyclic and modifying the amounts of the other elements in the proportions in which they occur in seawater. The assumption is somewhat arbitrary, but the resulting figures undoubtedly give a truer picture of the net contribution to the sea of material in solution.

A complete summary developed by Garrels and Mackenzie showing the estimated annual flux of various chemical species into the ocean basins is given in Table 9.8. Much of this material is in the form of suspended solids and does not immediately affect the chemical composition of the ocean water as a whole.

Evidently factors other than the supply of dissolved material in runoff from the land regulate the composition of seawater. Many reactions take place in the sea to alter the balance of dissolved substances. Adsorption and base exchange by particles of sediment remove some ions from solution; other ions react with sedimentary material to form new minerals, as for example glauconite and phillipsite. Biological activity is responsible for the extraction of much of the dissolved material. This is particularly true for calcium carbonate, which forms the hard parts of marine organisms, and for silica, which is used by diatoms and radiolaria.

It is not known for certain whether the addition of dissolved matter has brought about progressive changes in the relative composition of seawater during geological time. In any event, such changes must be exceedingly slow. The total quantity of dissolved solids in the ocean, assuming an average salinity of 35‰ and a mass of

Table 9.7   Composition of Dissolved Solids in River and Seawater

| | River Water | | Sea Water (%) | River Water less Cyclic Salts (%) |
	ppm	Percent		
HCO$_3$	58.5	48.6	0.41	54.6
SO$_4$	11.2	9.3	7.68	9.6
Cl	7.8	6.5	55.04	0.0
NO$_3$	1	0.8	—	0.9
Ca	15	12.5	1.15	13.9
Mg	4.1	3.4	3.69	3.4
Na	6.3	5.3	30.62	2.5
K	2.3	2.0	1.10	2.1
Fe	0.7	0.6	—	0.7
SiO$_2$	13.1	11.0	—	12.3
Sr, H$_3$BO$_3$, Br	—	—	0.31	—
	120.0	100.0	100.	100.0

Table 9.8 Estimated Annual Flux of Various Chemical Species into and out of the Ocean Basins (All Values in Units of 10^{14} g)

Chemical Species	Streams Dissolved	Streams Suspended	Ground Water	Aerosols	Dust	Marine Erosion	Ice	Subsurface	Totals	Percent of Total
SiO_2	4.4	121.0	0.5	—	0.4	1.5	13.0		140.8	60
Al	—	14.9	—	—	0.1	0.2	1.8		17.0	7
Fe	0.2	12.2	—	—		0.1	0.6		13.1	6
Ca	5.0	5.5	0.6	—		0.1	0.6		11.8	5
Mg	1.4	2.8	0.2	0.1		0.1	0.3		4.9	2
K	0.8	3.9	0.1	0.1		0.1	0.6		5.6	2
Na	2.1	2.0	0.2	0.7		—	0.5		5.5	2
Cl	2.6	—	0.3	1.4		—	—	0.1	4.4	2
SO_4	3.8	—	0.4	0.2		—	—		4.4	2
HCO_3	19.0	8.3	2.0	—		0.2	—	0.3	29.8	12
Total	39.3	170.6	4.3	2.5	0.5	2.3	17.4	0.4	237.3	
Percent of total	16.6	72.0	1.8	1.0	0.2	1.0	7.3	0.2		100

Others (NO_3^- organic matter, fluorine, sulfide, etc.) $\sim 3 - 4$
Estimated grand total $\sim 240 \times 10^{14}$ g/yr

After Garrels and Mackenzie, *Evolution of sedimentary rocks*, W. W. Norton, New York, 1971.

1413×10^{21} g, is 49.5×10^{21} g. The total amount of material in solution contributed annually by runoff from the land (42×10^{14} g) is only an infinitesimal fraction, 10^{-7}, of this amount; it is nevertheless significant in terms of the length of geological time.

THE BALANCE OF DISSOLVED MATTER IN SEAWATER

It is generally believed that the ocean is, to a large degree, in a steady-state condition, whereby the amount of an element introduced per unit time is balanced by an equal amount deposited in the sediments. Accepting this hypothesis of a steady-state condition, we can then define a residence time in years for each element, being the total amount of that element in solution in the oceans divided by the amount introduced by the rivers per year, corrected for airborne cyclic material.

$$T = \frac{A}{dA/dt}$$

Calculated residence times for many elements are given in Table 9.3.

Although many independent assumptions are involved in these calculations, both the absolute and relative values for most of the residence times seem reasonable. Sodium has the longest residence time, within an order of magnitude of the age of the oceans. This long residence time reflects a lack of reactivity of sodium in the marine environment; it is not readily incorporated in the common sedimentary minerals, nor is it removed by biological reactions. The abundant cations K, Ca, and Mg have residence times of the order of 10^6 to 10^7 years. Elements of low abundance in seawater have short residence times, especially those that might potentially be available in considerable amounts. The very low residence time of manganese, 1400 years,

is evidently related to its removal from solution by oxidation to the tetravalent state and precipitation as the manganese dioxide nodules that are abundant over wide areas of the ocean floor. Silicon and aluminum are among the elements with the shortest residence times. Much silicon is removed from solution by the activity of organisms; aluminum apparently has no biological function, but is probably rapidly removed from solution as newly precipitated clay minerals.

The significance of residence times for practical geochemical problems is illustrated in predicting the behavior of artificial radionuclides introduced into the ocean by nuclear explosions or as wastes from reactors. The radioiosotopes of strontium and cesium, both hazardous to living organisms, have very high residence times, relative to their half-lives. Hence, the radionuclides of these elements, if introduced into the ocean, are not rapidly precipitated; they remain in solution and are available to marine organisms.

Goldschmidt made an interesting study of the balance between supply and removal of several elements in seawater. The basis of his comparison was the total amounts of different elements supplied by weathering, disintegration, and sedimentation during geological time. According to Goldschmidt, about 160 kg of igneous rocks have been weathered for each square centimeter of the earth's surface. Since there are 278 kg of ocean water per square centimeter, approximately 600 g of rock have been weathered for each kilogram of water in the oceans. The 600 g of igneous rock have therefore been the potential source of the dissolved matter supplied by weathering to 1 kg of seawater. Of course, only a part of this 600 g of igneous rock has actually dissolved and remained in solution. Goldschmidt drew up a balance sheet relating the potential supply of several elements in 600 g of igneous rock to the amounts present in solution of 1 kg of seawater. The percentage in solution was obtained by dividing the amount of each element in seawater by the potential supply. The most striking feature is that the concentrations of a few elements are far greater than can have been supplied by weathering. These are common anions of seawater; chloride, sulfate, borate, and bromide. Either they were present in great amounts in the primeval ocean or they have been supplied largely from volcanic gases and thermal springs throughout geological time. Judging by the data available on amounts and composition of igneous exhalations, the latter alternative is adequate to explain at least the major part of these "superabundant" elements in sea water.

The concentration of a given element in seawater is a function of its rate of removal by chemical or physical means as well as its rate of addition. Mechanisms of removal include entrapment in pore spaces in sediments, chemical and biochemical precipitation, ion exchange reactions with clays and zeolites, and precipitation as evaporites in isolated basins. To evaluate the chemical environment in which these reactions may take place, a knowledge of the ionic concentrations, pH, and oxidation-reduction relationships in seawater is important. Such data are not always readily available and, hence, definite predictions of chemical reactions are often difficult. Most of the data in the chemical literature are for dilute solutions and only qualitatively applicable to seawater. In evaluating simple solubilities not only must changes in solubility products and ionic activities be taken into consideration but changes in ionic species as well. Table 9.3 shows that the chemical species for many

Table 9.9   Solubility of Some Trace Elements in Seawater

Element	Insoluble Compound	Calculated Concentration moles/L	Observed Concentration moles/L
Sr	$SrCO_3$	3×10^{-4}	9×10^{-5}
Ni	$Ni(OH)_2$	2×10^{-3}	1×10^{-7}
Co	$CoCO_3$	2×10^{-7}	7×10^{-9}
Ba	$BaSO_4$	1×10^{-7}	1.5×10^{-7}
La	$LaPO_4$	1×10^{-11}	2×10^{-11}
Th	$Th_3(PO_4)_4$	2×10^{-12}	1×10^{-12}

elements in seawater are not simple cations but complexes with the predominant anions. Table 9.9 illustrates the differences between expected and observed concentrations of elements in sea water based on their expected most insoluble salt. It may be seen that the sea is greatly undersaturated in all of these ions with the possible exception of barium, lanthanum, and thorium.

Krauskopf (1956) has made an extensive study of the factors controlling the concentrations of Zn, Cu, Pb, Cd, Ni, Co, Hg, Ag, Cr, Mo, W, and V in seawater. He investigated in detail four processes for the removal of these elements: precipitation of insoluble compounds with ions normally present in sea water; precipitation by sulfide ion in local regions of low oxidation potential; adsorption by materials such as ferrous sulfide, hydrated ferric oxide, hydrated manganese dioxide, and clay; and removal by the metabolic action of organisms. He showed both by calculation and experiment that sea water is greatly undersaturated in all these ions; that is, precipitation of insoluble compounds cannot be responsible for the observed concentrations. Local precipitation of sulfides is a possible control mechanism for some of the elements but is probably not the chief control because the concentrations are unrelated to sulfide solubilities. Adsorption is a possible mechanism for all elements except V, W, Ni, Co, and Cr; if Cr is assumed to be removed by local reduction and precipitation of the hydroxide and the other four elements by biological processes, the existing concentrations can be fairly adequately accounted for. Adsorption processes supplemented by biological removal also furnish an explanation for the distribution of minor and trace metals in marine sedimentary rocks.

With few exceptions therefore all the elements have been potentially available in much larger amounts than are actually present in seawater. The solubility of its compounds, physicochemical factors such as adsorption and coprecipitation, and biological activity control the removal of an element from solution. Probably some ions in river water are precipitated as insoluble compounds as soon as they enter the sea because of the difference in chemical environment. Some elements are removed from solution by complex reactions between ions in solution and solid material, which may be particles in suspension or bottom sediments. Such processes are termed halmyrolysis.

An interesting example of halmyrolysis is the accumulation of radioactive material in the upper layers of deep-sea sediments. The breakdown of ^{238}U in seawater results

in the production of ^{230}Th (ionium), which is coprecipitated almost quantitatively with ^{232}Th and accumulates in the bottom sediments. giving them abnormally high radioactivity. The half-life of ^{230}Th is 83,000 years; hence, the amount falls off with the age of the sediment; that is, the ^{230}Th to ^{232}Th ratio is greatest in the surface layers and decreases below. It thus is possible to estimate the age of different layers in deep-sea cores. Some valuable results on rates of deposition of deep-sea sediments have been obtained by this method; it has been shown that red clay in the central Pacific accumulates at a rate of about 1 mm in 1000 years, whereas the red clay in the Atlantic generally shows a deposition rate that is considerably more rapid.

A remarkable feature of seawater is the very low percentages of certain poisonous metals and metalloids, including some, such as selenium and arsenic, that form soluble complex anions and therefore might be expected to accumulate. The quantities of these elements potentially supplied to the ocean during geological time are so great that a serious poisoning action would have resulted had not some process been active in eliminating them.

Biological activity is undoubtedly of great importance in controlling the concentrations of many of the elements in the sea, not only the abundant ones, such as calcium and carbon, but also rarer ones, such as copper, vanadium, and phosphorus. For example, the concentration of copper and iron is extremely low in seawater, but one or the other is an essential constituent of the blood of many marine animals. Organisms remove from solution elements that would not otherwise precipitate and, if conditions are such that some of the organic material becomes a permanent part of the sediments, this process must play an important part in controlling seawater composition.

Of particular interest and importance in the chemistry of seawater are the reactions that control the pH and redox potential. The oceans function effectively as buffer systems having the ability to assimilate both acids and bases without major changes in pH. The pH of ocean water sampled near the surface almost always has a pH of between 8.1 and 8.3. The buffering action in seawater has been attributed to both carbonate-CO_2 and silicate systems. In the carbonate-CO_2 system the buffering reactions and related equilibrium constants at 25°C are

$$CO_2 \text{ (gas)} + H_2O \text{ (liquid)} = H_2CO_3 \text{ (aqueous)} \quad K_p = 10^{-1.5}$$
$$H_2CO_3 \text{ (aqueous)} = H^+ \text{ (aqueous)} + HCO_3^- \text{ (aqueous)} \quad K_1 = 10^{-6.4}$$
$$HCO_3 \text{ (aqueous)} = H^+ \text{ (aqueous)} + CO_3^= \text{ (aqueous)} \quad K_2 = 10^{-10.3}$$
$$CaCO_3 \text{ (solid)} + H_2O \text{ (liquid)} = Ca^{++} \text{ (aqueous)} + CO_3^= \text{ (aqueous)} \quad K_{SP} = 10^{-8.3}$$
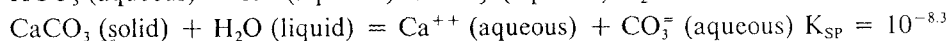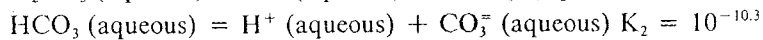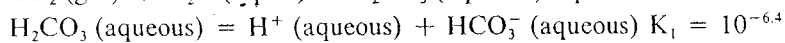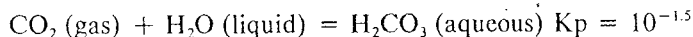
For an atmospheric CO_2 pressure of 3×10^{-4} atmospheres the equilibrium pH for the system is close to 8. This model assumes that the CO_2 pressure is independently fixed by biological activity. Large changes in this CO_2 partial pressure would result in only small changes in pH. Addition of acids through volcanic or other sources would lead to the dissolution of $CaCO_3$ and the eventual production of CO_2, again with little change in pH. Alternatively the reaction of silicates with seawater in reactions of the type

THE NATURE OF THE HYDROSPHERE

$$Na^+ \text{ feldspar} + H^+ \text{ (aqueous)} = \text{kaolinite} + Na^+ \text{ (aqueous)} + SiO_2$$
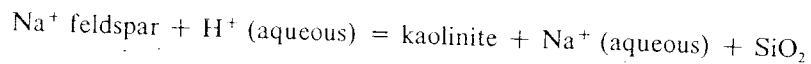
may fix the pH of the ocean and hence regulate the atmospheric pressure of CO_2. In addition the system $H^+ + H_2BO_3^- = H_3BO_3$ and other minor reactions may also contribute to the buffering system. This complex question has not been completely resolved due to the complexity of the ocean system. It may not be unreasonable to conclude that all reactions play a part; the carbonate-CO_2 system on a short time scale and the ion-exchange silicate system on a longer time scale. Mackenzie and Garrels (1966) suggested that reactions combining both systems of the type

$$\text{Al silicates} + \text{cations} + SiO_2 + HCO_3^- = \text{cation Al silicate} + CO_2 + H_2O$$
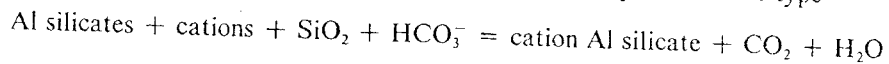
may play a large part in controlling the distribution of cations in seawater.

In normal seawater the half reaction that controls the oxidation potential (Eh) is

$$2H_2O = O_2 + 4H^+ + 4e^- \quad Eh \,(25°C) = 1.23v$$
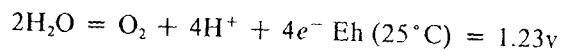
The concentration of dissolved molecular oxygen in most parts of the ocean is very close to saturation relative to the atmosphere so it is a highly oxidizing solution. The oxygen partial pressure may be increased by oxygen production by phytoplankton or reduced by organic decay. For a partial pressure of oxygen of 0.2 atm and a pH of 8 at 25°C the Eh of seawater is calculated to be 0.85v. Measured values are usually somewhat lower.

THE HISTORY AND EVOLUTION OF THE OCEAN

A review of the literature reveals that ideas concerning the history and evolution of the ocean are quite speculative. The early history of the hydrosphere, like that of the atmosphere, is tied to the loss of volatiles from the earth to its surface. Whether this process has been fairly constant over geologic time or occurred at a more rapid rate early in the earth's history is open to question. Also of interest are changes in the chemical composition of seawater. Possibilities include the addition of dissolved salts to water essentially fresh or the constant buildup of seawater with a composition similar to what we have today. Paleontological and stratigraphic records provide evidence that the dissolved salts in ocean water have not greatly differed over the last 600 million years and most likely over the last 3 billion years. Minor changes, such as the addition of oxygen to the atmosphere and hence a variation in the oxidation potential of the seawater, would not affect the mean dissolved salt content but would allow iron in the ferrous form to exist in the water. Under present conditions iron is oxidized to the ferric form and precipitated.

A consideration of this history of the oceans involves three main concerns: the source and rate of accumulation of water, the addition of anions, and the addition of cations.

The chemical evolution of the oceans may be discussed in terms of the following hypotheses.

1. The water condensed from a primeval atmosphere and the chloride has been added throughout geological time (constant volume-volcanic chloride hypothesis).
2. Both water and chloride are the result of initial condensation (constant volume-constant chloride hypothesis).
3. Both components have appeared gradually (volcanic ocean-volcanic chloride hypothesis).
4. The chloride was present as metallic chloride on the earth's surface, and the water has been added throughout geological time by volcanic activity (volcanic ocean-constant chloride hypothesis).

The uniformitarian view that the composition of the ocean has been similar to that at the present day as far back as the Archean is borne out by the available data. We know that magmas contain water vapor and gases such as carbon dioxide, hydrogen chloride, hydrogen fluoride, and hydrogen sulfide, and that, therefore, these have been continually added to the hydrosphere. Much of the water, and some of the other gases also, may be cyclical. On the whole, it appears that no one of the above theses can be accepted as presenting the full story. Current ideas on the origin of the earth suggest the formation of an ocean as soon as the surface was cool enough for water to condense. The amount of water in and the salinity of the primitive ocean are unknown. Throughout geological history the oceans have grown by the addition of primary magmatic water. This growth may well have been considerably more rapid in early geological time, when a thinner crust, higher average temperature, and a greater production of radiogenic heat probably resulted in more vigorous igneous activity. The nature of sedimentary rocks and fossils indicates that the content of dissolved substances has probably remained much the same, approximate constancy of composition being maintained by the complex chemical reactions outlined in the previous section.

Whether igneous activity has provided sufficient water to account for all or the greater part of the hydrosphere is still a matter of dispute. For example, Verhoogen (1946) concluded that the total lava extruded since the beginning of the Cambrian cannot exceed a volume of 3×10^{22} cm^3 or, adopting an average density of 3.3, a mass of 1×10^{23} g. Even if this lava were accompanied by as much as 10% of its mass of water, not more than 0.7% of the ocean can have been added since the beginning of the Palaeozoic or less than 5% during the whole of geological time. However, Verhoogen's estimate does not appear to include submarine volcanic activity, which is probably as widespread as terrestrial volcanism and would thus increase his figures by a factor of three.

If the oceans have increased in volume throughout geological time, we are faced with the problem of correlating this with the relationship between continents and ocean basins. If the amount of seawater has increased continually, then the ocean basins must have increased in area or depth or both. The areal relationship of continents and ocean basins during Precambrian time is very poorly known. Since for the most part the oldest rocks are concentrated in the central shield areas, with younger rocks around their edges, it has been suggested that continental crust has grown throughout geological time. This implies a considerable reduction in the area

occupied by the ocean, and the sea floor must necessarily have become deeper in relation to the continental level.

For a discussion of the chemical changes taking place in seawater, Garrels and Mackenzie (1971) have divided ocean history into three time periods: the earliest oceans, the oceans 1.5 to 3.5 billion years before the present, and modern oceans. They speculate that the earliest oceans would be made up of water with dissolved gases such as H_2S, HCl, HF, and HBr. The chemically active ocean would vigorously attack the earth's crust to dissolve out cations and silica, leaving a residue of alumina and silicates. This would neutralize the acidic anions in the water, creating an ocean perhaps with a salinity similar to that today. As the seawater became more neutral CO_2 would be dissolved as well.

Our oldest sedimentary records from 3.5 billion years ago support the uniform ocean idea. Between 3.5 and about 1.5 billion years ago there is evidence for slow but continuous chemical trends in the sedimentary rocks and the ocean composition. Dominant mineral-ocean interaction at that time would be of the type

$$14H_2O + 6NaAlSi_3O_8 + 3FeCO_3 + 15MgSiO_3 + 3CaCl_2 = 22SiO_2$$

$$\text{albite} \qquad \text{siderite} \qquad \text{enstatite} \qquad\qquad \text{silica}$$

$$+ 6NaCl + 3Mg_5Al_2Si_3O_{10}(OH)_8 + 3CaCO_3 + Fe_3Si_2O_5(OH)_4$$

$$\text{chlorite} \qquad\qquad \text{calcite} \qquad \text{greenalite}$$

$$5MgSiO_3 + CaAl_2Si_2O_8 + CO_2 = Mg_5Al_2Si_3O_{10}(OH)_8 + CaCO_3 + 4SiO_2$$

$$\text{enstatite} \qquad \text{anorthite} \qquad\qquad\qquad \text{chlorite} \qquad\qquad \text{silica}$$

and

$$Fe_3Si_2O_5(OH)_4 + 3CO_2 = 3FeCO_3 + 2SiO_2 + 2H_2O$$

$$\text{greenalite} \qquad\qquad \text{siderite} \quad \text{silica}$$
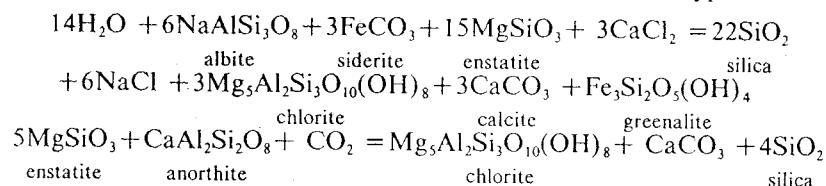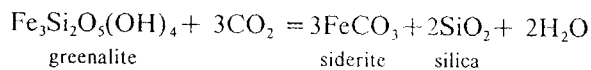
The unique characteristic of these reactions is the presence of ferrous iron. The reactions as written contain minerals that are important in the 2-billion-year-old banded iron formations. Likewise many of the rocks in the 3.5-billion-year-old Onverwaacht formation of South Africa are cherts, indicating a higher dissolved silica concentration in the early oceans. Today this is controlled by organic activity. The calculated CO_2 pressure at 25°C in an atmosphere coexisting with siderite, greenalite, and amorphous silica is about $10^{-2.5}$ atm—an order of magnitude greater than it is today. The pH would be about neutral. About 1.5 billion years ago the seawater began to take on its modern characteristics with a mean seawater composition in approximate equilibrium with calcite, K-feldspar, illite-montmorillonite, and chlorite. There is some evidence for periodic compositional changes but these are of only minor magnitude.

SELECTED REFERENCES

Baas-Becking, L. G. M., I. R. Kaplan, and O. Moore (1960). Limits of the natural environment in terms of pH and oxidation potentials. *J. Geol.* **68**, 243–284.

Broecker, W. S. (1974). *Chemical oceanography.* 214 pp. Harcourt Brace Jovanovich, New York. A general view of chemical oceanography with many references to more detailed studies.

Ellis, A. J. and W. A. J. Mahon (1977). *Chemistry and geothermal systems.* 392 pp. Academic Press, New York. A good review of hot waters and brines.

Garrels, R. M. and F. T. Mackenzie (1971). *Evolution of sedimentary rocks.* 397 pp. W. W. Norton & Co., New York. An interesting and useful approach to material transfer by water.

Holland, H. D. (1978). *The chemistry of the atmosphere and oceans.* 351 pp. John Wiley & Sons, New York.

Kitano, Y. (ed.) (1975). *Geochemistry of water.* 455 pp. Dowden, Hutcheson & Ross Inc., Stroudsburg, Pa. A collection of selected papers on water chemistry. Benchwark Papers in Geology, Volume 16.

Livingstone, D. A. (1963). *Chemical composition of rivers and lakes. U.S. Geological Survey Prof. Paper, 440 G.*

Martin, D. F. (1970). *Marine chemistry,* Vol. 2, 451 pp. Marcel Dekker, New York. Theory and applications of element cycles in the ocean. Volume 1 discusses analytical techniques.

Riley, J. P. and R. Chester (1975–76). *Chemical oceanography (second ed.).* Academic Press. New York. Six-volume work with 34 chapters by individual authors on the chemistry of ocean waters and their interaction with sediments.

Rubey, W. W. (1951). Geologic history of sea water. *Bull. Geol. Soc. Amer.* **62,** 1111–1147. A critical discussion of the evolution of the oceans, based on a careful evaluation of the geochemical data.

Stumm, W. and J. J. Morgan (1970). *Aquatic chemistry.* 583 pp. Wiley-Interscience, New York. An introduction to chemical equilibria in natural waters.

10 THE NATURE OF THE BIOSPHERE

The concept of the biosphere was introduced by Lamarck. The term is used in two senses, one denoting that part of the earth capable of sustaining life and the other the sum total of living matter—plants, animals, and microorganisms. In the first sense it appears to comprise a narrow zone at the surface of the hydrosphere and the lithosphere, and this indeed is the region in which life is most prolific, thanks to the favorable conjunction of water, air, and radiant energy. However, the zone of life is more extensive than this: insects and spores have been collected at high altitudes, living organisms have been dredged from the ocean bottom at great depths, and bacteria have been found in oil-well brines from within the upper crustal layers.

The influence of the biosphere in geochemistry covers an even greater range. Non-living materials of biochemical origin are accumulated or transported over large areas of the earth and greatly influence the distribution and accumulation of many elements. The geochemistry of the biosphere may be considered as biogeochemistry as in the Russian school developed by Vernadsky, or as organic geochemistry. In either case it is important to realize that it has an arbitrary distinction from bio-chemistry. Whereas biochemistry is primarily concerned with chemical reactions taking place within living organisms, organic geochemistry is concerned with the movement of elements controlled by living organisms. Such an elemental balance would entail the kinetics of the movement of elemental matter in and out of organisms as well as the mass of the biosphere both living and dead. Although the biosphere has a small mass, large amounts of material move through it in dynamic equilibrium and it has strong interactions with its surrounding environment. In many cases it is difficult to separate the geochemical role of living substances from the geochemical role of organic material formed after the destruction of the organisms.

The biosphere controls many of the earth's unique characteristics. Biogeochemical reactions often approach equilibrium only slowly; hence, nonthermodynamic-equilibrium situations are often sustained in the biosphere.

The biosphere developed later than the other geological spheres, since life can hardly have begun before surface conditions were much as they are today, presumably with an atmosphere and a hydrosphere. How much later is largely a matter of speculation. Well-preserved fossils first appear in Cambrian formations, but organisms complex enough to leave fossil remains must surely have been preceded by a long evolutionary cycle. Even in very ancient Archean formations, slates and schists containing free carbon are found, and many scientists have assumed that this free carbon is the residue of some form of life. It is of course also possible to synthesize organic materials by inorganic reactions, and it is perhaps not unreasonable to explain the initial accumulation of some organic compounds on the earth's surface by a large influx of organic-rich carbonaceous chondrites or comets early in the earth's history. The oldest sedimentary rocks containing carbon compounds are the Isua rocks in Greenland which are approximately 3.7 billion years old. Structures believed to be those of primitive algae have been recognized in limestones of the Bulawayan system of southern Zimbabwe, one of the oldest Precambrian systems, with an age of about 2.6 billion years. On the north shore of Lake Superior well-preserved microfossils identified as bacteria and blue-green algae are found in cherts of the Gunflint formation, about 1.9 billion years old. It thus seems probable that life began more than three billion years ago, but evolved comparatively slowly until the development of green plant photosynthesis, apparently about 1.8 to 2 billion years ago. Once the atmospheric oxygen had risen to sufficiently high levels for respirational metabolism (approximately one to three percent of present atmospheric oxygen), metazoa could develop. This was perhaps about one billion years ago.

In the material sense the biosphere is made up of a group of complex organic compounds. Living matter consists of a relatively small number of organic molecules including alcohols, fatty acids, amino acids, and purines—out of which are built up the more complex molecules of the carbohydrates, proteins, fats, nucleic acids, and lignin. Life has one common material characteristic, the presence of protein molecules, and one common physiochemical process, the stepwise formation of organic substances carried out practically isothermally. Living matter can, in general, be divided into the two great classes: plants, or producers, and animals, or consumers. Because animals are completely dependent on vegetable matter for their nourishment, they cannot be considered as independent organisms but rather as parasitic on the plants that constitute their food. The assimilation of carbon dioxide by means of chlorophyll and light is of dominant importance for life on this planet. The process can be expressed by the following equation:

$$6CO_2 + 6H_2O = C_6H_{12}O_6 + 6O_2$$
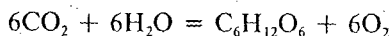
For the reaction as written $\Delta G = 688,000$ cal and $\Delta H = 673,000$ cal. Because ΔG is positive, energy must be supplied if the reaction is to go from left to right; this energy is provided by solar radiation. In this way the radiant energy of sunlight is converted into the chemical energy of organic compounds. The reverse reaction is

one of biological oxidation, or respiration, and the energy thus liberated may appear as heat or as work. A fairly close balance presumably exists between photosynthesis and respiration, although over the whole of geological time respiration has been exceeded by photosynthesis, the surplus energy derived therefrom being stored in preserved organic matter, in part as coal and petroleum. Photosynthesis cannot add oxygen to the earth's surface unless some organic products are withdrawn from reoxidation.

Microorganisms are critical to the geochemical reactions of the biosphere. The bulk of organic reactions take place through these minute but abundant organisms. Some chemical changes catalyzed by organisms, like photosynthesis, are unique, having no inorganic analogs. Others could only take place slowly or rarely in natural systems. Microorganisms allow the fixation of nitrate from molecular nitrogen, and in anaerobic systems derive energy by the reduction of sulfate to sulfide and the reduction of iron. Such reactions often leave their fingerprints on the geochemical cycle. The action of microorganisms can also control the chemical environment in limited areas. In unbuffered pools the rapid removal of CO_2 through photosynthesis can allow a rise in alkalinity to a pH of over 10 with a high Eh, whereas the action of thiobaccilli can in other environments lead to waters with a pH of less than 1. The presence of reducing compounds such as CH_4, phenols, and H_2S can create a very low Eh in biologically active areas.

Although most organic material formed in the biosphere is rapidly recycled to H_2O and CO_2, it plays an important role in geochemical reactions along the way. Particularly evident is the influence of organic materials in weathering reactions. Without organic acids and organic chelating agents, as well as the direct action of microorganisms on minerals, chemical weathering would take place much more slowly. Those organic molecules not immediately destroyed by the action of microorganisms can undergo progressive change by nonbiological processes. The natural biopolymers can be broken down into simpler compounds by hydrolysis. The more resistant complexes undergo diagenesis through burial and heating in sediments to form both monomers and higher polymers such as natural gas, petroleum, kerogen, and coal. Metamorphism can further degrade these materials to anthracite and graphite, with the further loss of simple hydrocarbons and carbon dioxide.

THE MASS OF THE BIOSPHERE

Looking on the biosphere as the sum total of living matter, we may attempt to determine its mass. Here, however, we are faced with many difficulties, difficulties that were not present in other geochemical calculations. The matter of the biosphere is not uniformly distributed, as is, for example, that of the atmosphere. It is also in a constant state of change, and the cycle of changes is very rapid. The life span of any organism is minute in comparison to geological time, and the life cycles of different organisms are immensely varied—compare, say, the life of a redwood tree with that of a protozoan.

In spite of all these difficulties, quantitative work in biogeochemistry demands some idea of the total amount of matter in the biosphere. Borchert (1951) estimated

that the total amount of carbon in living organisms is 2.8×10^{17} g. The total annual productivity is also an important figure in discussions of the biosphere. Recent estimates for the total annual production of organic carbon are 118×10^9 tons for terrestrial environments and 55×10^9 tons for marine environments, a grand total for the earth of 173×10^9 tons. As we will see in the section on the carbon cycle, the mass of carbon in living material is several thousands times less than the nonliving organic or graphitic carbon in crustal rocks.

The mass of the biosphere is insignificant in comparison to that of other geochemical zones. According to Rankama and Sahama, the relative weights of the hydrosphere, the atmosphere, and the biosphere are expressed by the following figures:

Hydrosphere	69,100
Atmosphere	300
Biosphere	1

In spite of the negligible mass of the biosphere, however, it is a zone of great chemical activity, and its geochemical effects are of considerable significance. Its significance can best be realized in terms of the turnover of material taking place therein. If the mass of the biosphere has been approximately constant over the last 500 million years and if the average life cycle were one year, then the total amount of matter that has passed through the biosphere in that time is comparable with the total mass of the earth. Of course, most organisms appear to us as inconspicuous blobs of organic matter that have a certain life cycle, reproduce, die, and leave their surroundings very much as they were before. This picture is, however, a false one. We noted in a previous chapter how most of the free oxygen of the atmosphere has been produced by the chemical activity of plants. An individual organism secreting calcium carbonate does not change the character of the earth much, but during geological time such organisms have taken calcium in enormous amounts from a dilute solution in seawater and thereby formed most of the carbonate rocks. An even more remarkable separation is the deposition of silica by radiolaria and diatoms from seawater, which contains as little as 0.02 to 4g/silicon per ton. Millions of square kilometers of the ocean floor are covered with siliceous sediments, and cherts, representing deposits of such organisms, are common in the geological column. The organically deposited carbon in sediments has been extracted from an atmosphere which throughout geological time has probably never contained more than a few hundredths of a percent of carbon dioxide. In addition, every organism has some metal compounds in its structure, and the biosphere is thus of considerable significance in the migration and concentration of many elements.

In considering the effect of the biosphere on geochemistry it appears that much of the change from living material takes place (immediately after death) by the decomposing action of microorganisms, followed by the deposition of organic material and the subsequent dispersal or concentration of the chemical elements by the preserved organic compounds. Questions commonly considered by organic geochemists include whether an organic or carbon-rich material is biogenic or abiogenic in origin, whether it was originally terrestrial or marine, and whether the present material was formed *in situ* or has migrated through permeable materials.

THE COMPOSITION OF THE BIOSPHERE

If an estimate of the mass of the biosphere is difficult, an estimate of its average composition is even more so. Living matter varies greatly in composition. The analytical data are rather variable in quality, and our knowledge of the relative amounts of different organisms present in the biosphere are little more than educated guesses. Water is the principal component of all organic matter; wood contains about 50% and vertebrates about 66%, whereas marine invertebrates may be more than 90% H_2O. Table 10.1 provides data on the ultimate composition of biological material. The metabolic role of the invariable elements is fairly well understood. They can be roughly classified into energy elements (C, O, H. N), macronutrients (P, Ca, Mg, K, S, Na, Cl — Na is essential for animals, but not for plants), and micronutrients (Fe, Cu, Mn, Zn, for both animals and plants; B, Mo, and Si for plants only; and Co and I for animals only). A few of the variable elements, such as vanadium and bromine, may play an important play in the metabolism of some organisms, but for many of them no biological function is known. Many trace elements, whether or not they are essential, are noxious above quite low levels. Differences in the composition of various species or individuals exist, but the similarities are on the whole greater than the differences.

In all, over 60 elements have now been recorded in analyses of organisms, but many of these are "ultramicro" constituents. Great difficulty is experienced in cleaning organisms for analysis, since inorganic matter on the surface and in the digestive system is not easy to remove, and many published data are therefore suspect. The statement that an element is "present" in one species and "absent" from another

Table 10.1 Distribution of Elements as Percentage Body Weight of Organisms

Invariable			Variable		
Primary 60–1	Secondary 1–0.05	Microconstituents <0.05	Secondary	Microconstituents	Contaminants
H	Na	B	Ti	Li	He
C	Mg	Fe	V	Be	Ar
N	S	Si	Br	Al	Se
O	Cl	Mn		Cr	Au
P	K	Cu		F	Hg
	Ca	I		Ni	Bi
		Co		Ge	Tl
		Mo		As	
		Zn		Rb	
				Sr	
				Ag	
				Cd	
				Sn	
				Cs	
				Ba	
				Pb	
				Ra	

may give unwarranted prominence to comparatively small differences in concentration that happen to span the limits of sensitivity of the analytical method used. For an element to have biogeochemical significance it must either be shown to participate in the life of the organism or else to be consistently present in concentrations greater or much smaller than that in the environment.

When the data on the occurrence of the different elements in organisms are plotted on the periodic table, certain relations are apparent. The invariable elements (except iodine and molybdenum) are all of low atomic number, belonging to the first four horizontal rows in the periodic table.

The probability that an element plays an important part in metabolism depends on its ability to participate in the aqueous colloidal systems constituting the bodies of living organisms. Common elements are more likely to occur than rare ones, elements usually soluble in aqueous systems more than those highly insoluble. Hutchinson (1943) showed that ionic potential is significant in this. He showed that elements of comparable crustal abundance are very unequally taken up by plants. When the ratio between the percentage of an element in plant material and the average in rocks is plotted on Goldschmidt's ionic potential diagram (Figure 10.1), the correlation is readily seen. Elements of low ionic potential, which form soluble cations, and those of high ionic potential, which form soluble anions, are readily assimilated. The elements of intermediate ionic potential which form insoluble hydrolysates are unavailable and generally play no part in metabolism. Since plants sustain herbivorous animals and, ultimately, carnivorous animals as well, the same features apply to these organisms.

Since plants make up the bulk of living material and animals depend on plants for sustenance, an estimate of the distribution of elements in marine and land plants may give a first approximation of the chemical composition of the biosphere. Such a compilation put together by Bowen (1966) is given in Table 10.2. Note that the concentrations are based on dry plant tissue and hence do not include the hydrogen and oxygen from the water in the analyzed organisms. Also included in Table 10.2 are the estimated abundances in soils. The average soil is a mixture of mineral matter with living and dead organic material.

Some data on chemical composition of specific organisms are reproduced in Table 10.3. If we compare these figures with those for the average abundances of the same elements in the earth's crust, we find that a few are decidedly enriched in organisms. Hydrogen and oxygen head the list, of course, since organisms are so largely water, but carbon, nitrogen, phosphorous, and sulfur are also concentrated by the vital activities of animals and plants. Considered in terms of the ash, instead of the living material, several more elements are enriched—Na, K, Br, and I in a marine invertebrate such as *Calanus finmarchicus;* Ca, K, Cl, Mg, and B in alfalfa; P, Na, K, Cl, Br, and I in humans. On the average, however, plants and animals concentrate relatively few of the elements. Exceptions are known, especially for plants, that are often capable of taking up large amounts of an element if it is readily available in the soil. A classic example is afforded by some of the plants associated with seleniferous soils of the western Great Plains of the United States; selenium contents as high as 1.5% have been reported in some specimens of locoweed *(Astragalus racemosus)*. The ability of certain plants to reflect the presence of unusual concentrations
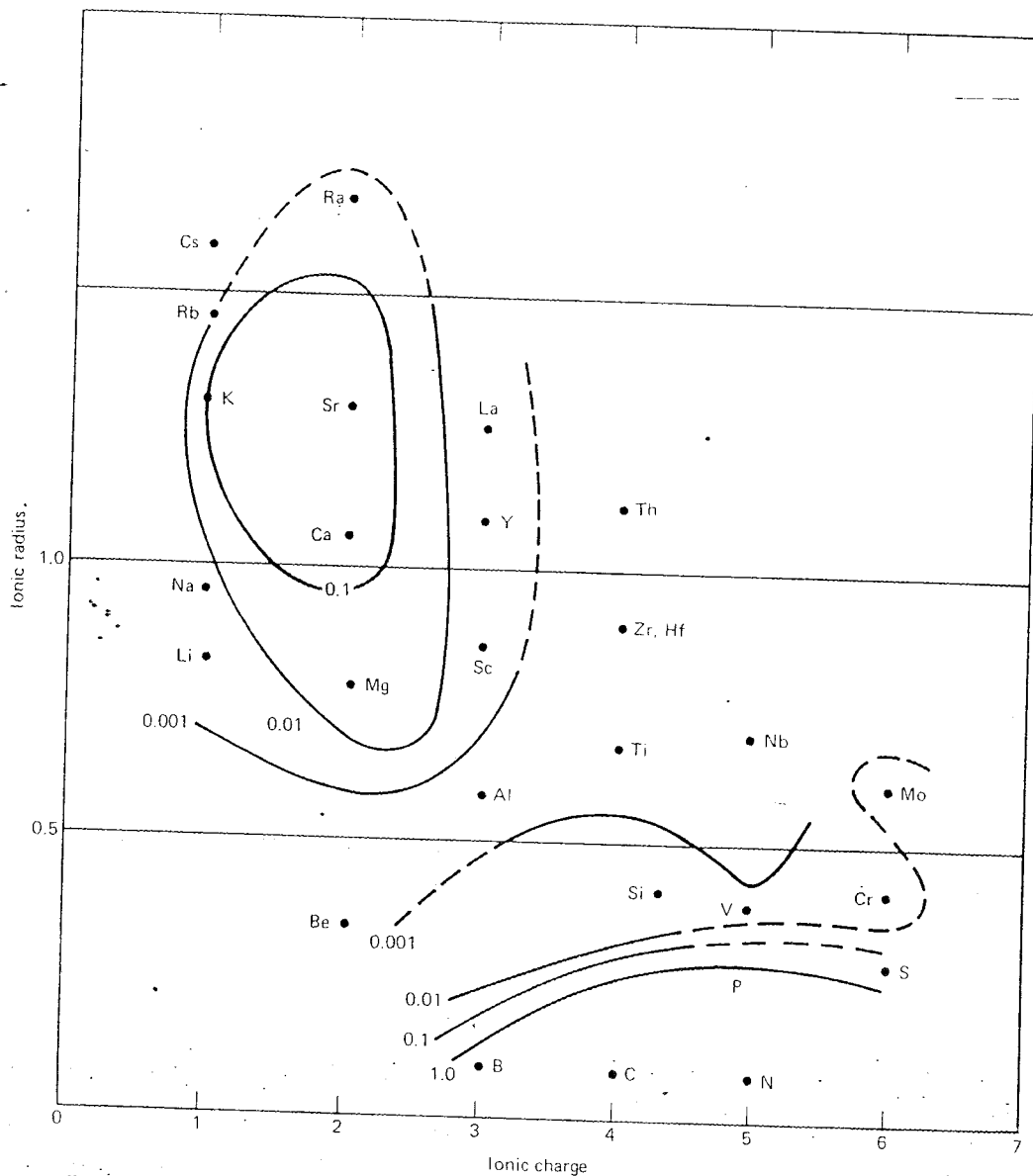
Figure 10.1   Concentration ratio of elements in terrestrial plants and in the crust, in relation to ionic potential. (Hutchinson, 1943, p. 337)

of a particular element in the soil and subsoil has been applied to the search for hidden ore bodies. Biogeochemical prospecting is based on the accumulation of trace elements in plants, whereas geobotanical prospecting concentrates on specific plants that grow where a specific trace element is present in the soil.

Generally, however, the most remarkable examples of chemical concentration through biological processes are provided by marine organisms. The ability of many

Table 10.2   Distribution of Selected
Elements in Dry Tissue of
Marine and Land Plants. (in
ppm)

Element	Marine Plants	Land Plants	Soils
Ag	0.25	0.06	0.1
Al	60	500	71,000
As	30	0.2	6
Au	0.012	0.002	—
B	120	50	10
Ba	30	14	500
Br	740	15	5
C	345,000	454,000	20,000
Ca	10,000	18,000	13,700
Cd	0.4	0.6	0.06
Cl	4700	2000	100
Co	0.7	0.5	8
Cr	1	0.23	100
Cs	0.07	0.2	6
Cu	11	14	20
F	4.5	0.5	200
Fe	700	140	38,000
Ga	0.5	0.06	30
Hg	0.03	0.015	0.03
I	1500	0.42	5
K	52,000	14,000	14,000
La	10	0.085	30
Li	5	0.1	30
Mg	5200	3200	5000
Mn	53	630	850
Mo	045	0.9	2
N	15,000	30,000	1000
Na	33,000	1200	6300
Ni	3	3	40
O	470,000	410,000	490,000
P	3500	2300	650
Pb	8.4	2.7	10
Rb	7.4	20	100
S	12,000	3400	700
Se	0.8	0.2	0.2
Si	1500	200	330,000
Sr	1400	26	300
Ti	12	1	5000
V	.2	1.6	100
W	0.035	0.07	1
Zn	150	100	50
Zr	20	0.64	300

After HMJ Bowen, *Trace elements in biochemistry.* Academic Press, 1966.

Table 10.3 Elementary Composition (Weight Percent) of Different Organisms

Calanus Finmarchicus (a copepod)*			Alfalfa (lucerne)†			Man‡		
O	79.99		O	77.90		O	62.81	
H	10.21		C	11.34		C	19.37	
C	6.10	n. 10^0	H	8.72	n. 10^0	H	9.31	n. 10^0
N	1.52		N	8.25		N	5.14	
Cl	1.05					Ca	1.38	
Na	5.4	n. 10^{-1}	P	7.06	n. 10^{-1}	S	6.4	n. 10^{-1}
K	2.9		Ca	5.80		P	6.3	
S	1.4		K	1.70		Na	2.6	
P	1.3		S	1.037		K	2.2	
						Cl	1.8	
Ca	4	n. 10^{-2}	Mg	8.2	n. 10^{-2}	Mg	4	n. 10^{-3}
Mg	3		Cl	7.0				
Fe	7	n. 10^{-3}	Si	9.3	n. 10^{-3}	Fe	5	n. 10^{-3}
Si	7		Fe	2.7		Si	4	
B	1.5		Al	2.5		Zn	2.5	
Br	9	n. 10^{-4}	B	7.0	n. 10^{-4}	Rb§	9	n. 10^{-4}
I	2		Rb	4.6		Cu	4	
Cu	6		Mn	3.6		Sr	4	
			Zn	3.5		Br	2	
			Cu	2.5		Sn	2	
			F	1.5		Mn	1	
			Mo‖	1		I	1	
V	4	n. 10^{-5}	Ti	9	n. 10^{-5}	Al	5	n. 10^{-5}
Mo	2.4		Ni	5		Pb	5	
Ti	2		Br	5		Ba	3	
			Ti	4.6		Mo§	2	
			V	1.6		B	2	
Co	6	n. 10^{-6}	I	2.5	n. 10^{-6}	As	5	n. 10^{-6}
			Co	2		Co	4	
						Li§	3	
						V§	2.6	
						Ni	2.5	

As,Sn,Pb,Sr,Ba n. 10^{-n}

*Vernadsky, Z. *Krist. Mineral. Petrog. Abr. B. Mineral Petrog. Mitt.* **44**, 191, 1933.
†Water, 75.1%; org., 22.4; ash, 2.45%; from Bertrand, 1950, p. 442.
‡Water, 60%; org., 35.7%; ash, 4.3%; from Bertrand, 1950, p. 442.
§Mean figure for mammals.
‖ Normally in plants, 3.10^{-5}.

marine invertebrates to extract calcium carbonate and of diatoms and radiolaria to extract silica has already been mentioned. Bromine is concentrated as dibromoindigo, the classical Tyrian purple dye, in some species of *Murex*. Some species of the marine ascidians effect the remarkable concentration of vanadium to 0.01%, compared to its concentration of 0.0000002% in seawater.

Some sponges and corals are rich in iodine, but marine plants are particularly effective in extracting this element, so much so that certain seaweeds were once harvested to provide commercial iodine. Oysters commonly build up a concentration of copper 200 times that in seawater; along with many other marine animals their respiratory fluid is hemocyanin, a close relative of hemoglobin, but with copper instead of iron as an essential constituent.

THE BIOGENIC DEPOSITS

We have been considering the composition of living organisms as a background to the understanding of deposits in the geological column resulting from biological activity. Such deposits may be called bioliths and can be divided into two groups: the noncombustibles and combustibles. The most important noncombustible is limestone, which is discussed in Chapter 6. Siliceous deposits of organic origin are much less widespread. The biological origin of nonclastic siliceous rocks in the geological column is difficult to prove, since the silica in organisms (diatoms, radiolaria, and sponges) is present as opal, which is easily reconstituted by diagenesis with the obliteration of organic structures. However, sea water is not saturated with SiO_2, and therefore the initial deposition of silica from solution in sea water must be biogenic. Chert is a general term for sedimentary nonclastic siliceous material, and it may occur dispersed as nodules, generally in limestone, and sometimes as the major part of whole geological formations hundreds and occasionally thousands of feet thick. Bramlette (1946) has given a careful account of one of these great chert formations, the Monterey of California, and summarized the work of previous investigators on this and similar deposits. The Monterey chert lends itself to the elucidation of the mode of formation of these rocks, since it is geologically young (Miocene) and little altered. Its original nature is therefore more certainly deciphered than that of older formations in which diagenesis and mild metamorphism have caused solution and redeposition of the silica with the obliteration of primary structures. Bramlette concludes that the silica was deposited originally as organic remains, mainly diatoms. In effect, the Monterey cherts and probably most other such deposits are primarily organic in origin.

Microorganisms probably play a considerable role in the formation of many biogenic deposits. The evidence for this is carefully presented by Beerstecher (1954). Bacteria and algae are capable of great chemical activity. Several types precipitate hydrated ferric oxide, utilizing the energy of a reaction of the type

$$4FeCO_3 + O_2 + 6H_2O = 4Fe(OH)_3 + 4CO_2 + 40,000 \text{ cal}$$
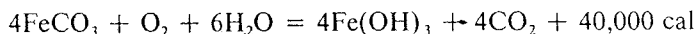
Some geologists consider that organisms have been responsible for the deposition of iron ores either by this type of mechanism or the indirect oxidation of ferrous iron by oxygen-producing algae. Living organisms can also be the mechanism by which

extensive marine sedimentary phosphate deposits may be formed. Insoluble calcium phosphate from the organism's hard parts and phosphate in the organic molecules may both contribute to deposits of calcium phosphate.

Because of the delicacy of microorganisms they seldom leave any recognizable trace behind them, so the positive identification of fossil bacteria is exceedingly difficult. However, the demonstration of a bacterial origin for deposits in the geological column may be obtained indirectly. A case in point is the occurrence of sulfur deposits in sedimentary rocks. These have been investigated by Feely and Kulp (1957). Anaerobic bacteria can reduce sulfates to hydrogen sulfide, and the formation of these sulfur deposits has been ascribed to their activity. This hypothesis has been greatly strengthened by work on the sulfur isotopes. Bacterial reduction of sulfate to sulfide results in a higher $^{32}S/^{34}S$ ratio in the sulfide than the sulfate from which it originated. Sulfide and sulfur of the deposits in salt domes of Louisiana and Texas have higher $^{32}S/^{34}S$ ratios than the associated sulfate, indicating enrichment in ^{32}S during their formation. It is believed that bacterial reduction of calcium sulfate produced hydrogen sulfide, which then reacted with more calcium sulfate to produce the sulfur of these deposits. In this case the bacteria use the sulfate as their energy source and petroleum hydrocarbons as their source of carbon-rich building materials.

Sulfides are among the most ubiquitous biogenic minerals. Sulfate is reduced by bacteria to form hydrogen sulfide, which reacts with iron to form pyrite. Pyrite, a common constituent in sedimentary rocks, may act as a scavenger to accumulate chalcophilic trace elements.

The great majority of the organic material in the earth's crust is in the form of dispersed molecules in the sedimentary rocks. Figure 10.2 illustrates the carbon cycle, and shows that dispersed material greatly outweighs the concentrated end members of the two principal divisions, coal and petroleum. Geochemically, all these materials represent concentrations of carbon compounds produced by organic activity. The study of the origin of the biogenic material presents the following problems: the nature and composition of the parent organisms; the mode of accumulation of the organic material; and the reactions whereby it was transformed into the end products.

As indicated earlier, microorganisms play a large part in the degradation and decomposition of biogenic molecules. Depending on the types of molecules and the chemical environment, individual biogenic molecules may be preserved, degraded, or completely decomposed. The initial biopolymers in living organisms first undergo hydrolysis and microbial degradation, setting free simpler biomonomers such as amino acids, carbohydrates, fatty acids, and other organic molecules. Some of these molecules are destroyed by metabolism or oxidation in place or after removal by solution. Some, like the pigments or porphyrins, are very resistant chemically and can be preserved in the sediments. With sedimentation and polymerization, residual and new geopolymers may be accumulated by inorganic processes. These would include "humic acid" complexes, polymerized lipids, and resins. This material may be further changed by diagenesis and thermal reactions to form still more complex polymers and geomonomers cracked off by the thermal reactions. Products of these reactions include petroleum, natural gas, bitumen, kerogen, and coal. Further metamorphism may form carbon-rich anthracite and graphite, together with the mobile
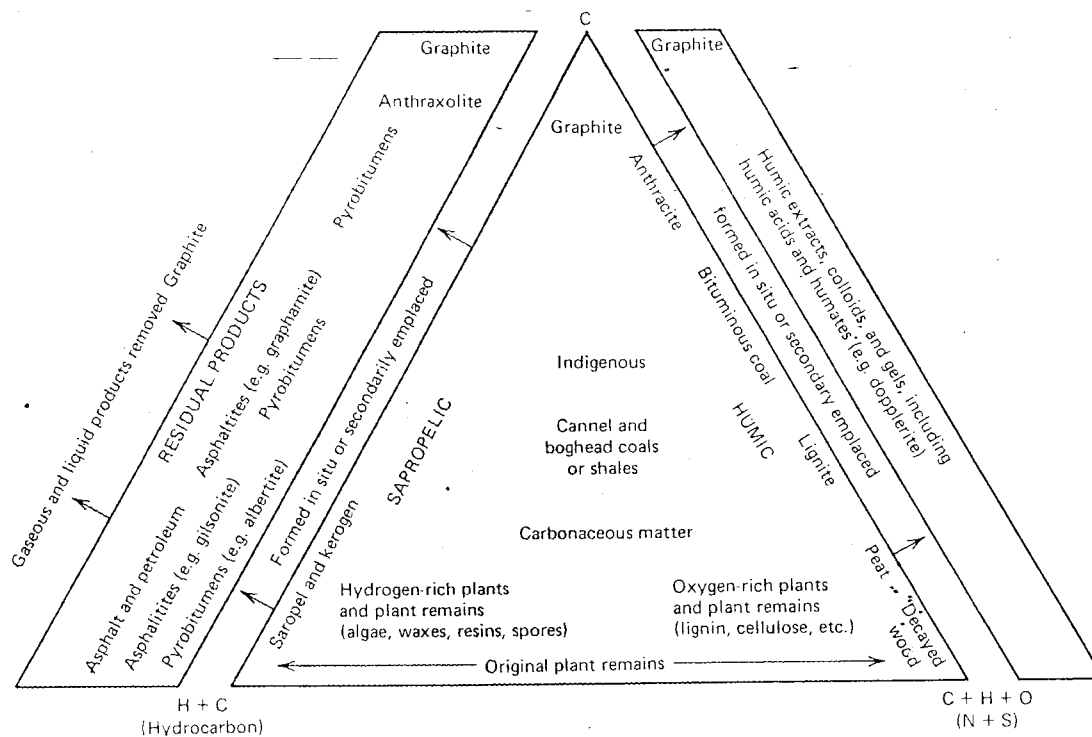
Figure 10.2 Diagram showing generalized relation among the types of carbonaceous materials. (After Devine et al., The role of humic acids in the geochemistry of uranium. *Proc. 2nd Intern. Confr. Peaceful Uses of Atomic Energy*, vol. 2, 1958).

light molecules, carbon monoxide, carbon dioxide and methane. Figure 10.2 illustrates the geochemical relationship of the various types of carbonaceous materials. Primary starting materials may be of two types, sapropelic or humic. Humic materials, the remains of woody plants, contain significant amounts of oxygen, nitrogen, and sulfur in addition to their carbon and hydrogen, and are the precursors of coal. Sapropelic materials are the remains of fresh water or marine lower-order plants, and are the precursors of petroleum. Kerogen is a somewhat vaguely defined term for dispersed carbonaceous material in sedimentary rocks. Pyrobitumen is carbonaceous material insoluble in CS_2; bitumen is the soluble fraction. The proportion of pyrobitumen to bitumen increases as sapropelic materials undergo progressive diagenesis.

THE ORIGIN OF COAL

Geological evidence provides ample justification for the belief that coal has been formed from terrestrial plant remains. The chief constituents of plant material are celluose and lignin, although a great variety of minor components—proteins, essen-

tial oils, organic acids and their salts, tannins, etc.—enter into the composition of vegetable matter. The empirical formula of cellulose is $C_6H_{10}O_5$ and that of lignin approximately $C_{12}H_{18}O_9$, so that they do not differ greatly in the relative proportions of carbon, hydrogen, and oxygen. The structure of lignin is, however, aromatic—that is, made up of ring-like groups of carbon atoms, the benzene ring being a familiar example—in contrast to the aliphatic nature of cellulose. On this account some authorities believe that lignin is the chief mother substance of coal, since the breakdown products of coal are largely aromatic. In addition, lignin is resistant to attack by microorganisms, whereas cellulose is readily decomposed into carbon dioxide, methane, and aliphatic acids. However, it has been shown that cellulose can be converted into aromatic compounds under conditions similar to those believed to exist during coal formation. The story is, of course, more involved than the simple conversion of lignin or cellulose into coal substances. Neither cellulose nor lignin contains nitrogen, yet nitrogen compounds are important in coal. Some of these may be derived from minor constitutents, such as plant proteins, but the nitrogen content of coal is relatively high compared with that of most plants. It has been suggested, therefore, that bacteria play a major part in the first stages of the formation of coal, since the nitrogen content of bacteria runs up to 13%. Oxygen in coal is considered to be present as hydroxyl (or phenolic), carboxyl, methoxyl and carbonyl groups. This is illustrated in Figure 10.3, which shows that in the higher rank coals the phenolic and carbonyl groups are retained but the others have been depleted.
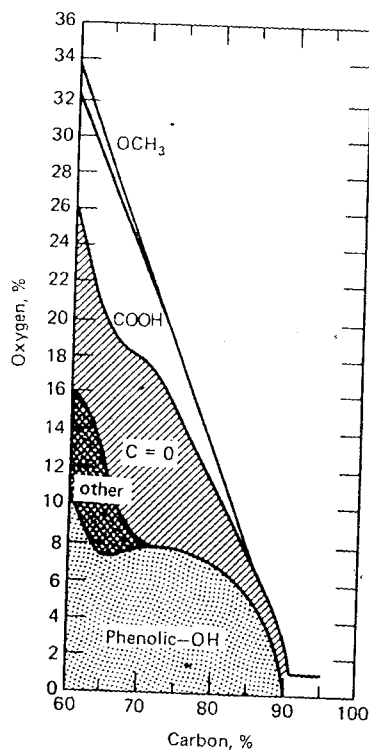
Figure 10.3 Functional groups with oxygen in various coals. (After Van Krevelen, 1961.)

Table 10.4 The Average Composition of Fuels

	C	H	N	O
Wood	49.65	6.23	0.92	43.20
Peat	55.44	6.28	1.72	36.56
Lignite	72.95	5.24	1.31	20.50
Bituminous coal	84.24	5.55	1.52	8.69
Anthracite	93.50	2.81	0.97	2.72

From Clarke, *The Data of Geochemistry*, p. 773.

The transformation of vegetable matter to coal involves two stages, one largely biochemical and the other metamorphic. The biochemical stage is essentially part of the conversion of vegetable matter into peat, during which microorganisms are active in reconstituting the organic matter. Peat formation involves the following processes: (a) rapid decomposition of water-soluble substances; (b) slow decomposition of the cellulose compounds, the nature of the peat depending to a large extent on the completeness of decomposition of these compounds; and (c) gradual accumulation of the cell substance of the microorganisms, which accounts for the increase in nitrogen content. The biochemical stage is eventually brought to an end when conditions become unsuitable for bacterial activity, either by burial under inorganic sediments or by the development of toxic conditions within the organic matter. Beyond that stage metamorphism, the action of heat and pressure, is responsible for coalification.

The progressive change from wood through peat to coal can readily be followed by chemical analyses. Table 10.4 gives the average compositions of wood, peat, and three successively higher ranks of coal—lignite, bituminous coal, and anthracite. Rank is a measure of the degree of metamorphism a coal has undergone. Various suggestions have been made regarding the geological factors that have determined the rank of a coal. Some of these are (a) the length of time since burial (Paleozoic coals are generally high-rank, Tertiary coals low-rank), but many exceptions are known, and it is now believed that time has no direct influence on rank; (b) the action of heat from earth movements or from igneous intrusions; (c) pressure resulting from compression during folding and faulting; and (d) increased pressure and temperature resulting from depth of burial. Of these the latter is most readily evaluated and has been formulated into a rule (Hilt's rule): In a series of coal seams the fixed carbon increases and the volatile matter decreases with depth. Hilt's rule has been applied to many coal basins, and the decrease in volatile matter is generally of the order of 0.2–0.8% per 100 feet of descent. Suggate has analyzed the depth-volatile relationship in a number of coal-fields and has demonstrated (Figure 10.4) a general pattern of depth-volatile distribution for British coals from low-rank bituminous (42% volatile matter) to anthracite (5% volatile matter). He presents good evidence that depth of burial is the cause of rank increase, except in areas of contact metamorphism. For the transformation of peat into low-rank bituminous coal he estimates a depth of burial of 7500 feet and for transformation into anthracite with 5% volatile matter a depth of burial of about 19,000 feet. These data can be usefully applied in reverse to determine original depth of burial of coal seams whose present position is the result of tectonic movements or erosion.
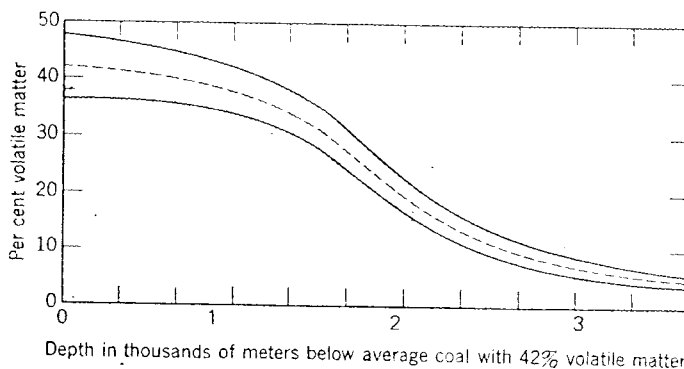
Figure 10.4   Depth-volatile relationship for British bituminous
and anthracite coals, calculated on a dry, ash-free
basis. (After Suggate, *Geol. Mag.* **93**, 212, 1956.)

Chemically, the passage from wood to anthracite is mainly an increase in carbon coupled with a decrease in oxygen. Hydrogen also decreases, but much less rapidly. The process is essentially one of reduction. The H : O ratio (weight percent), which is 1 : 8 in cellulose and 1 : 7 in wood, increases to 1 : 1 in anthracite. The changes in composition from peat to anthracite in terms of the changes in carbon and hydrogen are shown diagrammatically in Figure 10.5.

The processes outlined above give an empirical account of the transformation of plant material to coal. Little, however, is known of the chemical reactions involved. Van Krevelen (1961) has provided a useful summary of the state of knowledge in
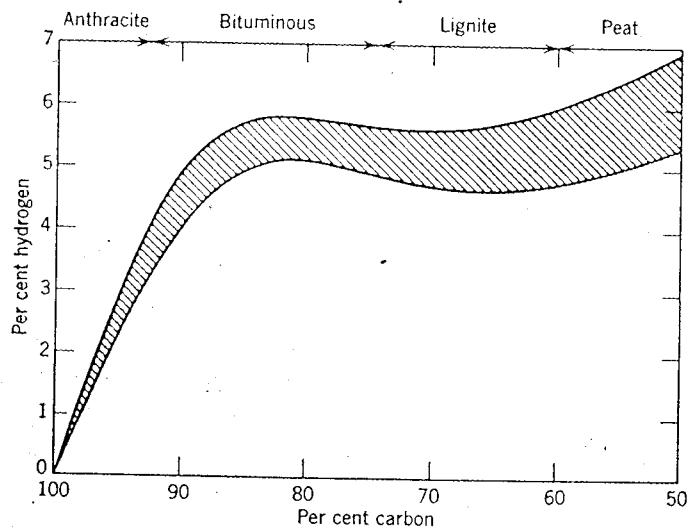
Figure 10.5   Compositional changes, calculated on a dry, ash-free
basis in the series from peat to anthracite, in terms
of the content of carbon and hydrogen the-so-called
coal band.

this important field. Research on the chemical constitution of coal has indicated that the coal material consists essentially of complex organic substances of high molecular weight. Study of the breakdown products of coal shows that these compounds are largely aromatic and that the degree of aromatization increases with rank, until ultimately complete aromatization to graphite may be reached. The oxygen is evidently present as OH and COOH groups and is expelled during rank increase as water and carbon dioxide, or in the initial stages by selective leaching by ground water. The rapid drop in hydrogen content from bituminous coal to anthracite is a reflection of the expulsion of methane at this state of metamorphism. Nitrogen is present in coal as amino groups or substituting for carbon in the ring structures. The evidence of the chemical reactions of coal suggests that the carbon in the ring systems is largely saturated; that is, the rings are naphthenic, and that six-membered rings predominate.

Maximum alteration of coaly material produces graphite; this alteration is analogous with the metamorphism of rocks, although the organic compounds in coal react more readily to increasing temperatures and pressures than do minerals. Coal composition is in fact an extremely sensitive indicator of the degree of alteration not only of its own material but also of the enclosing rocks. However, even at the anthracite stage the enclosing rocks are usually indurated but would not be described as metamorphic. In a few areas, such as the Narragansett basin of Rhode Island (Quinn and Glass, 1958), it has been possible to follow in the field and in the laboratory the transition from coal metamorphism to rock metamorphism. The change in the coal is indicated by an increasing degree of graphitization, as shown by the appearance of graphite reflections in an X-ray diffraction pattern. The anthracite is practically amorphous to X-rays, and the enclosing rocks show only a minor degree of recrystallization. With increasing metamorphism the rocks grade successively into chlorite schists and then into biotite schists. The meta-anthracite in the chlorite schist gives fairly well-defined graphite reflections, and in the biotite schist it is a well-crystallized graphite.

THE ORIGIN OF PETROLEUM

General agreement exists that the source material of petroleum was biological and was laid down in a sedimentary environment, although the theory of the inorganic origin of petroleum dies hard. There is some uncertainty whether nonmarine sediments are to be entirely excluded as petroleum source beds, but the predominance of marine sediments in oil-bearing formations supports the belief that the parent material of most petroleum was deposited in a near-shore marine environment.

Important clues to the origin of petroleum may be derived from a knowledge of its composition. Crude oil shows a remarkable constancy in elementary composition. Generally it has a carbon content of 83 to 87%, a hydrogen content of 11 to 14%, and other elements, mainly oxygen, nitrogen, and sulfur, up to a maximum of about 5%. Over 99% of many crude oils consists of carbon and hydrogen only. The ash content is uniformly low, from 0.001 to 0.05%.

In contrast to this uniformity and simplicity of elementary composition, crude oil

shows a great variability and complexity with respect to the compounds that may be present. It consists of a variable mixture of a large number of different hydrocarbons. These hydrocarbons may be grouped into three homologous series: (a) paraffins, with the general formula C_nH_{2n+2}; (b) cycloparaffins or naphthenes, ring-type hydrocarbons with the general formula C_nH_{2n}; and (c) benzenoid hydrocarbons, with the general formula C_nH_{2n-6}. The proportions of these major types vary considerably from one crude to another, and crudes are distinguished as paraffinic, naphthenic, or benzenoid according to the predominant type.

If crude petroleum is assumed to be a mixture in thermodynamic equilibrium, it should be possible to calculate its composition, provided adequate thermodynamic data on the many possible hydrocarbons are available. Much work on the free energies of hydrocarbons has been carried out at the U.S. Bureau of Standards and elsewhere in recent years. Some of the significant results of this work are summed up in the following statements:

1. Paraffins are comparatively the most stable hydrocarbons at lower temperatures.
2. In an homologous series the stability increases as the number of carbon atoms decreases. Methane is the most stable hydrocarbon.
3. The stability of the naphthenes does not differ greatly from that of the paraffins.
4. In the temperature range most directly concerned with oil genesis benzenoid hydrocarbons are less stable than the corresponding saturated hydrocarbons.

Because of the wide variation in amounts of the different hydrocarbons in crudes of approximately the same bulk composition it is evident that crude oil is not a thermodynamic equilibrium mixture. The composition of a crude probably depends largely on the relative reaction rates of the competing reactions that have been responsible for its formation. However, oils from older strata as a rule contain more paraffins and more volatile (i.e., low molecular weight) hydrocarbons than those from younger formations. This suggests that increasing age is reflected in a closer approach to thermodynamic equilibrium.

Other clues as to origin provided by the composition of crude oil are the presence of heat-sensitive compounds, such as porphyrins and complex nitrogen-substituted hydrocarbons, which are decomposed at temperatures around 200°C. Temperatures of 200°C are not exceeded in petroleum formation, and probably the maximum temperature has generally been much lower. This is supported by an overall correlation between the occurrence of oil and the rank of coal within the same general area; oil cannot be expected when the rank of coal exceeds a certain value, best expressed by the fixed carbon content (fixed carbon in excess of 70% usually signifies the absence of petroleum). Oil is clearly more sensitive to metamorphism and is readily destroyed by conditions that merely cause an increase of rank in the coal.

The nature of the original organic matter from which crude oil has been formed is a subject of considerable controversy. It is conceivable that practically any organism, animal or plant, may have contributed to the formation of petroleum. Since, however, the plankton are the most abundant organic material in the sea and form the food of higher organisms, it is reasonable to consider their composition as representative of the composition of the raw material from which petroleum has been

Table 10.5 Elementary Chemical
              Composition of Organic
              Substances

Object	C	H	O	N
Phytoplankton	45	7	45	3
Zooplankton	50	8	32	10
Pelagic deposits	56	8	30	6
Continental plants	54	6	37	2.8
Humus of the soil	56	4.5	36	3.5

formed. Data on the chemical composition of plankton and other organisms as summarized by Vinogradov (1953) are given in Table 10.5. Clearly the conversion of the original material to petroleum must involve the efficient removal of oxygen.

Plankton, like higher organisms, contain carbohydrates, proteins, and fats, and small amounts of hydrocarbons have been found in diatoms. On death planktonic organisms are either eaten or sink to the bottom to become part of the marine sediments. Oceanographic investigations have revealed the environment of deposition whereby organic matter may be preserved for burial and ultimate transformation into petroleum.

Reducing conditions are clearly necessary, since oxidation rapidly decomposes organic matter into carbon dioxide and water. The environment must also be inimical to carrion-eating animals for which dead organisms serve as food. These conditions are present on the ocean floor in places where the bottom waters are stagnant, especially in basins and troughs. They are also present to a less extensive degree in swamps, constricted lagoons, and brackish marshlands. In such areas organic matter sinking to the bottom is not destroyed in the usual manner but is decomposed by the action of anaerobic bacteria. Putrefaction takes place, and the product is a black mud known as sapropel. A present-day example is the Black Sea, where water circulation is restricted and bottom sediments with as much as 35% organic matter have been recorded (the average marine sediment contains about 2.5% organic matter).

Sapropel is believed to be the parent substance of petroleum, just as peat is the parent substance of coal. The conversion of sapropel to petroleum probably involves both biochemical and inorganic process, although the relative part each of these plays is the subject of much difference of opinion. There is good evidence that bacterial action is important in converting carbohydrates and proteins into compounds that may change into hydrocarbons. Chemically, the conversion of organic matter to hydrocarbons can be represented schematically by the equation

$$(CH_2O)_n = xCO_2 + yCH_z$$

carbohydrate   carbon dioxide   hydrocarbons
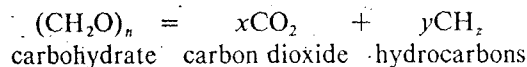
that is, a reciprocal oxidation and reduction in which part of the organic material is completely oxidized to CO_2 and part completely reduced to hydrocarbons. However, this is certainly an oversimplification. Biochemical processes probably first produce

fatty acids, which are later converted into hydrocarbons by a series of reactions involving fission. condensation. cyclation, and dehydration. These reactions may be promoted by catalysts, in which connection the significance of clay deserves to be emphasized. Clay particles are strong adsorbents and thereby bring foreign molecules into close contact. In addition, such molecules are not fixed at random but are in definite position relative not only to the clay, but also to each other, where they may be able to interact and form new compounds.

A controversial point of much significance in discussing the origin of oil is the time factor. Since little oil has been produced from rocks younger than Pliocene (and that may have migrated from older rocks), the widely accepted view has been that the formation of oil is a slow process. This view was seemingly confirmed by the failure of past investigators to detect liquid hydrocarbons in recent sediments. However, the situation was completely changed by the work of Smith (1954) and others, who have succeeded in extracting paraffin, naphthene, and benzenoid hydrocarbons from recent marine sediments. The quantity of free hydrocarbons varied up to 11,700 parts per million of dried sediment. That this material has formed in place and is not petroleum migrating from older formations is proved by carbon-14 dating, which has indicated an apparent age of about 10,000 years for the hydrocarbons. Extrapolation of the data for sediments of the Gulf of Mexico off Louisiana and the Santa Cruz Basin off California gives estimates varying from 4,500,000 to 10,400,000 barrels of crude oil per cubic mile of sediments in these areas. The problem of the ultimate source of petroleum thus appears solved; however, the hydrocarbons in recent sediments still differ significantly from the assemblages in crude oil. For example, the n-paraffin hydrocarbons extracted from recent sediments show a strong predominance of molecules with an odd number of carbon atoms, whereas no such preference is noted in crude oils. Hydrocarbons of low molecular weight are lacking in recent sediments, but they are abundant in many crude oils. Evidently the organic matter deposited with the sediments must undergo further evolution during the generation and accumulation of petroleum.

THE CONCENTRATION OF MINOR ELEMENTS IN BIOGENIC DEPOSITS

In terms of the average composition of the earth's crust, coal and petroleum are greatly enriched in carbon and hydrogen and sometimes in nitrogen and sulfur. In addition, the biogenic deposits often show remarkable concentrations of rare elements. The classic example was Goldschmidt's discovery of 1.6% GeO_2 in the ash of a coal from the Newcastle district in England (a figure that has been surpassed by the finding of up to 7.5% germanium in the ash of lignite from the District of Columbia). Goldschmidt followed up this discovery by a systematic study of the geochemistry of coal ash. Table 10.6 is a summary of some of his results. The figures in Table 10.6 are for the average content of "rich" ashes, and individual samples, such as the Newcastle ash with 1.6% GeO_2, may show much higher enrichment factors. Additional data are available in a paper by Swaine, in which he compares the trace element contents in New South Wales coal ash with those recorded in the literature (Figure 10.6). The remarkable feature of these results is the heterogeneous group of

Table 10.6   Rare Elements in Coal Ash

Element	Average Content in Coal Ash (g/ton)	Average Content in earth's crust (g/ton)	Factor of Enrichment
B	600	10	60
Ge	500	1.5	330
As	500	2	250
Bi	20	0.2	100
Be	45	2.8	16
Co	300	25	12
Ni	700	75	9
Cd	5	0.2	25
Pb	100	13	8
Ag	2	0.1	20
Sc	60	22	3
Ga	100	15	7
Mo	50	1.5	30
U	400	2.7	150
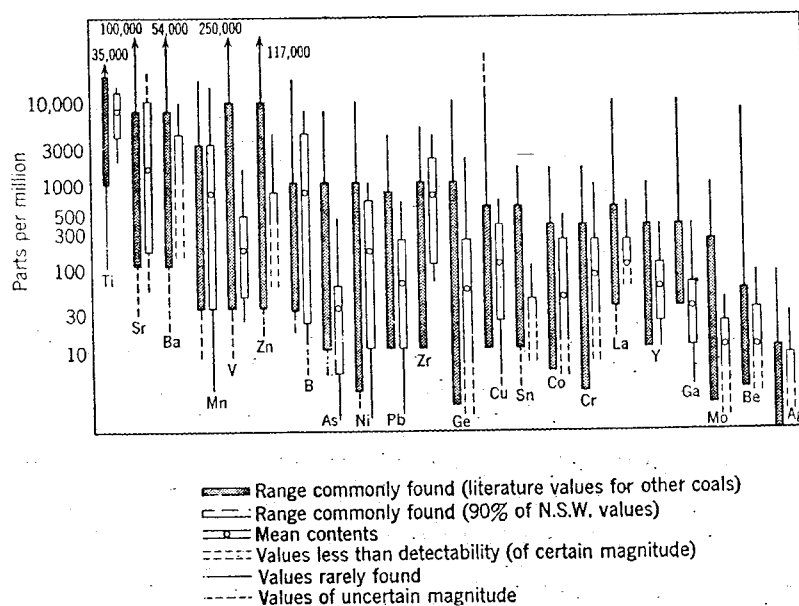
Figure 10.6   Range of trace element contents of coal ash; solid bars are for data from the literature, open bars are for determinations on New South Wales coals. (Swaine, *Technical Communication* **45**, Division of Coal Research, C.S.I.R.O., Australia, 1962.)

elements found in coal ash. These elements differ widely in their geochemical behavior. Both chalcophile and lithophile elements are present, although the chalcophile predominate. Ionic or atomic radii are apparently not significant, for the association in the coal ashes embraces elements with small, medium, and large radii.

Several possibilities may be suggested for the enrichment of a rare element in coal ash:

1. The element was accumulated by the vital processes of the plants that formed the parent material of the coal (possibly followed by selective decay whereby the rare element was further concentrated).
2. The element was precipitated from ground water by adsorption or chemical reaction during coalification.
3. The element formed part of the mineral matter deposited along with the organic material.

The third possibility can be discounted, because there is no reason to expect mineral matter deposited with organic material to be especially rich in rare elements. In addition, Goldschmidt's analyses showed clearly that the rare elements are not present in the mineral matter, since the concentrations of these elements are usually greatest in low-ash coals and least in those with high ash content. The first possibility may be responsible for some instances of enrichment of trace elements, although in general plants show little capacity to accumulate elements from the soil. Goldschmidt favored a concentration during decay of the plant remains whereby the more soluble elements were leached out, leaving others retained either as insoluble compounds or as metal-organic complexes. The second possibility, however, seems to provide particularly favorable circumstances for concentration of rare elements. The low oxidation potential of the medium would reduce sulfur compounds to H_2S, and thereby lead to the precipitation of chalcophile elements, many of which are markedly enriched in coal ashes. Another significant process is chelation of metallic ions by complex organic molecules, especially by phenolic groups present in humus in peats.

Experiments have shown that humus has considerable capacity for absorbing a variety of cations from solution and retaining them in stable combination. This process has been held responsible for the enrichment of uranium in some lignites, in which the uranium has been leached from overlying strata and is especially concentrated in the upper part of the lignite. Differences in individual trace elements in coals can then be attributed to the availability and concentrations of these elements in the groundwaters in contact with the low-grade stage of the coal. Subsequent loss of the oxygen-rich groups on the coal molecules apparently leaves the complexed trace elements behind to accumulate.

Compared with coal, petroleum (including asphaltites and bitumens) is far more specific in the concentration of rare elements. Vanadium shows a strong affinity for petroleum, and over 70% V_2O_5 has been recorded from petroleum ash. Some shipping companies have found it profitable to buy fuel oil from a particular locality, since the ashes can be sold as vanadium ore. In the famous Minasragra deposit in Peru the vanadium occurs with asphaltite. Goldschmidt believed that the typical elements

associated with petroleum and bitumen—vanadium, molybdenum, and nickel—are present as organometallic compounds which migrate with the hydrocarbons. He pointed out that these elements are effective catalysts for the synthesis of hydrocarbons; hence they may have been active in facilitating the formation of petroleum from organic remains. These elements may have been extracted from sea water by organisms that utilized them in the form of metal-organic porphyrin compounds. The porphyrin compounds are exceedingly stable and have been recognized in shales, asphalts, and petroleum dating back to the Paleozoic; they are evidently able to withstand the ordinary processes of diagenesis. The great enrichment of some of the trace elements indicates that a secondary exchange mechanism is most likely the reason for their concentration. Porphyrin molecules, which are themselves very stable, appear to exchange their original cations for others during the migration of the host petroleum in contact with connate marine and other groundwaters. If this is the case trace element concentrations give little help in learning the original source of the petroleum.

As discussed in Chapter 6, the black bituminous shales also show unusual concentrations of minor elements. An economically important example is the Mansfeld "Kupferschiefer" of Germany, which is a bituminous shale with a considerable copper content. It is worked as a copper ore and is markedly enriched in As, Ag, Zn, Cd, Pb, V, Mo, Sb, Bi, Au, and the platinum metals. The increased interest in uranium in recent years has revealed that this element, too, is often enriched in the black shales, contents of up to 400 g/ton being not uncommon.

THE GEOCHEMICAL CYCLE OF CARBON

Carbon, although not one of the more abundant elements in the earth, plays an important role, perhaps the most important one, in geochemistry, because carbon compounds are essential for every known form of life. The geochemistry of carbon is closely linked with that of the other essential elements of organisms, especially hydrogen, oxygen, nitrogen, and sulfur. Various aspects of the carbon cycle have been studied for over a century. In 1933, Goldschmidt developed the concept of the carbon cycle in both its biological and geological aspects and made quantitative estimates of the amount of carbon in the different parts of the cycle. Some of these estimates have been revised by later workers, especially Borchert (1951) and Wickman (1956). The geochemical inventory of carbon is given in Table 10.7. The great mass of it is tied up in rocks and does not become involved in the short-range cycle of carbon at the earth's surface. It should be remembered however that this buried carbon is equivalent to the free oxygen produced during geologic time.

Figure 10.7 illustrates the interrelation between the carbon cycle and the sulfur cycle and their mode of controlling the oxygen cycle.

Much work has been done in recent years on variations in the relative abundances of the stable isotopes of carbon, hydrogen, oxygen, sulfur, and nitrogen in or associated with organic material in the biosphere. Living organisms are capable of fractionating the isotopes of all these light elements. Generally they will enrich the lighter isotope of a pair, either in their cell tissues or in the metabolic products, as a

Table 10.7   Geochemical Inventory of Carbon in the Earth's Crust (Hunt, 1972)

	Graphitic or Organic Carbon*	Carbon as Carbon Dioxide or Carbonates*
Volcanic rocks	2,200	2,700
Granitic rocks	1,300	1,000
Metamorphic rocks	3,500	2,600
Sedimentary rocks		
Claystones and shales	8,900	9,300
Carbonates	1,800	51,000
Sandstones	1,300	3,900
Petroleum		
Nonreservoir	200	—
Reservoir	1	—
Coal	15	—
Dissolved in hydrosphere	0.7	38.3
Atmosphere	—	0.64
Biosphere (land and water)	0.3	
	~ 19,000	~ 71,000

*in units of 10^{18} grams.

function of kinetic reactions in which the lighter isotope reacts more rapidly. Craig (1953), Wickman (1956), and others have made many measurements of the $^{12}C/^{13}C$ ratio in materials from different geochemical environments and found small but regular variations (Figure 7.6). The results indicate that photosynthetic organisms are enriched by 10 to 20% in ^{12}C relative to the starting atmospheric or dissolved CO_2 ratios. Land plants show a greater ^{12}C enrichment than marine organisms reflecting a difference in $\delta^{13}C$ between atmospheric CO_2 of -7 to -10 and ocean HCO_3^- of -1 to -2. Degradation of the organic matter causes the regenerated CO_2 to be further enriched in the lighter isotope. As might be expected, repeated cycles of organic productivity and decay in a closed system may lead to enrichments of ^{12}C up to 9%. In organisms the light isotopes of H and C are preferentially enriched in fats. The high concentration of ^{12}C in petroleum is attributed to its origin from the degradation of fatty acids.

In 1941 Wickman, in one of the earliest applications of isotopic variations in geology, used the data then available on $^{12}C/^{13}C$ ratios in an ingenious calculation to determine the amount of organic carbon in sedimentary rocks. The basis for his calculation is that the differences in $^{12}C/^{13}C$ ratios between carbonate carbon and organic carbon represent fractionation of the two isotopes from original crustal carbon of uniform isotopic composition. If the total amount of carbonate carbon is A, organic carbon, B, their isotope ratios, x and y, and the isotope ratio of original crustal carbon, z, then

$$\frac{A}{B} \cong \frac{y - z}{z - x}$$

Wickman discussed the concept of "crustal carbon" at some length and has shown that the $^{12}C/^{13}C$ ratio for diamonds, for carbonates of probable magmatic origin, and
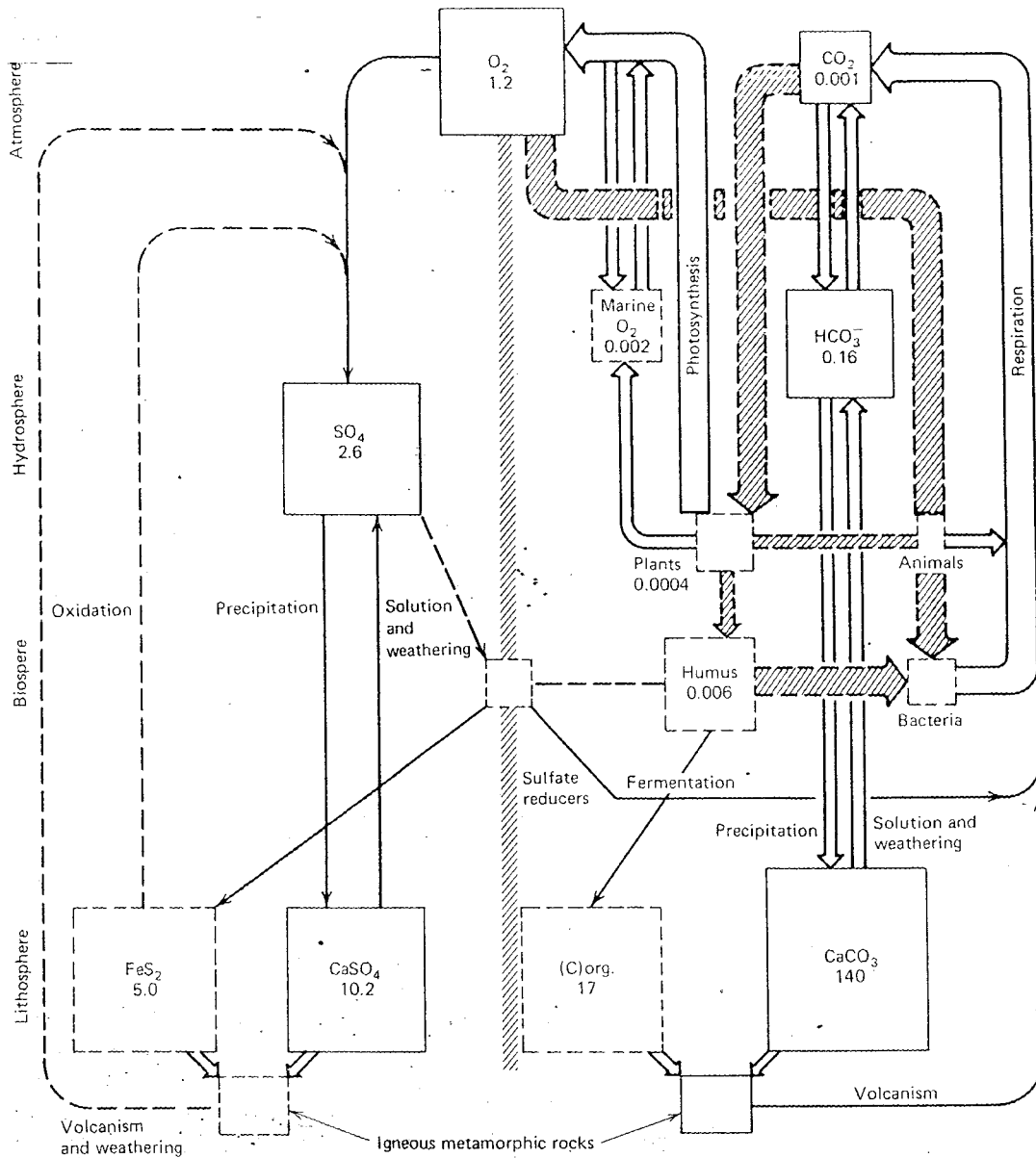
Figure 10.7 Interaction between the biogeochemical sulfur and carbon cycles and their mode of controlling the oxygen cycle. Line thickness denotes the approximate magnitude of annual flow rates. Numerical values signify total oxygen contents in 10^{15} metric tons either as molecular oxygen, oxygen sink (FeS_2), or equivalent amount of oxygen released (C organic).From R. R. Brookes and I. R. Kaplan in *Encyclopedia of Geochemistry* (R. W. Fairbridge, ed.), 1972.

for a number of graphites all give consistent values for this ratio around an average of 89.11, which he accepted as the figure for z. Then, using the figures given above for the $^{12}C/^{13}C$ ratio in carbonate rocks and in fossil organic carbon and his earlier estimate of 2420 ± 560 g/cm^2 of the earth's surface for A, he calculated that B is 700 ± 200 g/cm^2. This corresponds to a content of $0.41 \pm 0.13\%$ C in the average shale, a figure consistent with direct analytical data.

SELECTED REFERENCES

Abelson, P. H. (1978). Organic matter in the earth's crust. *Ann. Rev. Earth Planet. Sci.* **6**, 325–51. A good review of the nature of organic materials in the earth's crust.

Beerstecher, E. (1954). *Petroleum microbiology*, 375 pp. Elsevier Publishing, New York. This book is not only an exhaustive account of the role of bacteria in the origin and utilization of petroleum, but also includes an excellent chapter on the role of microorganisms in geological processes.

Borchert, H. (1951). Zur Geochemie des Kohlenstoffs. *Geochim. Cosmochim. Acta* **2**, 62–75. A detailed discussion of the carbon cycle, with quantitative data for the amounts of carbon in various parts of the cycle.

Bowen, H. M. J. (1966). *Trace elements in biochemistry*. 241 pp. Academic Press, London. A good source of information on the cycles of all elements in living materials and pollution problems.

Bramlette, M. N. (1946). The Monterey formation of California and the origin of its siliceous rocks. *U.S. Geol. Survey, Profess. Paper 212.* 57 pp. A detailed description of a thick and extensive nonclastic siliceous deposit and a discussion of the origin of this and similar formations. The author presents good evidence for the organic deposition of much of the silica.

Breger, I. A. (ed.) (1964). *Organic geochemistry.* 658 pp. Pergamon Press, New York. Contains 15 chapters by different authors covering the origin, classification, and geochemistry of the major biogenic deposits; excellent bibliographies.

Cannon, H. L. (1960). Botanical prospecting for ore deposits. *Science,* **132**, 591–598. A useful review of this subject, with an extensive bibliography.

Cloud, P. (1972). A working model of the primitive Earth. *Amer. J. Sci.* **272**, 537–548. A discussion of the interaction of the biosphere with the atmosphere, hydrosphere, and sedimentation.

————. (1976). Beginnings of biospheric evolution and their biogeochemical consequences. *Paleobiology* **2**, 351–387.

Colombo, U. and G. D. Hobson (eds.) (1964). *Advances in organic geochemistry.* 488 pp. Pergamon Press, New York. A symposium covering developments in the geochemistry of biogenic deposits.

Craig, H. (1953). The geochemistry of the stable carbon isotopes. *Geochim. Cosmochim. Acta* **3**, 53–92. A survey of the variation of the $^{12}C/^{13}C$ ratio in nature, based on measurements on several hundred samples of carbon from different geological environments.

Eglinton, G. and M. T. J. Murphy (eds.) (1969). Organic geochemistry—methods and results 828pp. Springer-Verlag, New York. A well-balanced source book of organic geochemical information including analytical techniques.

Feely, H. W. and J. L. Kulp (1957). Origin of Gulf Coast salt-dome sulfur deposits. *Bull. Amer. Assn. Petrol. Geol.* **41**, 1802–1853. These studies provide good evidence for the origin of native sulfur by the bacterial reduction of sulfate.

Hedberg, H. D. (1964). Geologic aspects of origin of petroleum. *Bull. Amer. Assn. Petrol. Geol.* **48**, 1755–1803. A critical review of the information from different approaches on the origin of petroleum; extensive bibliography.

Hunt, J. M. (1972). Distribution of carbon in the crust of Earth. *Bull. Amer. Assoc. Petrol. Geol.* **56**, 2273–2277. An inventory of carbon in the earth's crust.

Hutchinson, G. E. (1943). The biogeochemistry of aluminum and of certain related elements. *Quart. Rev. Biol.* **18**, 1–29, 129–153, 242–262, 331–363. A discussion of biological aspects of geochemistry and an account of the factors involved in the accumulation of elements through the vital activities of plants and animals.

———. (1954) The biogeochemistry of the terrestrial atmosphere. In *The earth as a planet.* G. P. Kuiper (ed.) (pp. 371–433). University of Chicago Press, Chicago. A critical review of the factors that influence the composition of the atmosphere, with special reference to biological processes.

Kenyon, D. H. and G. Steinman (1969). *Biochemical predestination.* 301 pp. McGraw-Hill, New York. A good review of the synthesis of life and biochemistry related to geological processes.

Kvenvolden, K. A. (1974). *Geochemistry and the origin of life.* 422 pp. Volume 14 in the Benchwork Papers in Geology Series. Dowden, Hutchinson and Ross, Stroudsburg, Pa.

Philippi, G. T. (1965). On the depth, time and mechanism of petroleum generation. *Geochim. Cosmochim. Acta* **29**, 1021–1049. One of a series of excellent papers on the geochemistry of petroleum.

——— (1974). The influence of marine and terrestrial source material on the composition of petroleum. *Geochim. Cosmochim. Acta,* **38**, 947–966. A discussion of variation in the nature of petroleum.

Quinn, A. W. and H. D. Glass (1958). Rank of coal and metamorphic grade of rocks of the Narragansett basin of Rhode Island. *Econ. Geol.* **53**, 563–574. An account of the chemical and mineralogical changes accompanying the transformation of coal into graphite.

Scientific American (1970). *The biosphere.* W. H. Freeman, San Francisco, 134 pp. A collection of articles on the basic fundemental biochemical cycles.

Smith, P. V. (1954). Studies on origin of petroleum: occurrence of hydrocarbons in recent sediments. *Bull. Amer. Assoc. Petrol. Geol.* **38**, 377–404. A classic paper describing the research that provided the first proof of the formation of liquid hydrocarbons in recent marine sediments.

Tissot, B. and F. Bienner (1973). *Advances in organic geochemistry 1973.* 1080 pp. Editions Technip, Paris. The sixth of a series of proceedings volumes from international meetings on organic geochemistry.

Van Krevelen, D. W. (1961). *Coal.* 514 pp. Elsevier Publishing, New York. A comprehensive account, with special reference to the origin. metamorphism, and chemical constitution of coal.

Vinogradov, A. P. (1953). *The elementary chemical composition of marine organisms.* 647 pp. Sears Foundation for Marine Research, Yale University, New Haven. An exhaustive compilation of all available data on the composition of marine organisims, originally published in Russian in three parts in 1935, 1937, and 1944.

11 METAMORPHISM AS A GEOCHEMICAL PROCESS

Metamorphism may be defined as the sum of the processes that, working below the zone of weathering, cause the recrystallization of rock material. During metamorphism, the rocks remain essentially solid; if remelting takes place a magma is produced, and metamorphism has passed into magmatism. Metamorphism is induced in solid rocks as a result of pronounced changes in temperature, pressure, and chemical environment. These changes affect the physical and chemical stability of a mineral assemblage, and metamorphism results from the effort to establish a new equilibrium. In this way the constituents of a rock are changed to minerals that are more stable under the new conditions, and these minerals may arrange themselves with the production of structures that are more suited to the new environment. Metamorphism accordingly results in the partial or complete recrystallization of a rock, with the production of new structures and new minerals.

It is not feasible to define a sharp boundary between diagenesis—the complex of processes that converts a newly deposited sediment into an indurated rock—and the onset of metamorphism. As was discussed in the preceding chapter, the alteration of peat to anthracite and ultimately to graphite is essentially a metamorphic process, but the enclosing sedimentary rocks show little change except induration until the carbonaceous matter has been considerably graphitized. Nearly all salt deposits show evidence of partial or complete recrystallization following burial, because they contain numerous minerals that are highly sensitive to the moderate increase in temperature accompanying burial to the depth of a few thousand feet. Many tuffaceous sediments, especially those that contained large amounts of volcanic glass, have been partly or completely recrystallized to zeolitic rocks, whereas interbedded lava flows are likely to show little alteration. Clearly, the onset of metamorphism cannot be

defined in terms of a fixed temperature and pressure, and depends also on the mineralogical composition and fabric of the rock. Even rocks of practically the same chemical composition, say a granite and a rhyolite ash, may react very differently, due in part to smaller grain size, higher permeability, and hence to a greater surface area exposed to fluids in the case of the ash. Metamorphism can be understood on the basis of the fundamental principles of thermodynamics and kinetics but the great variety of possible rock types and environments make predictions difficult.

The three factors of heat, pressure, and chemically active fluids are the impelling forces in metamorphism. The heat may result from the general increase of temperature with depth or to contiguous magmas. Pressures may be resolved into two kinds: hydrostatic or uniform pressure, which leads to change in volume; and directed pressure or shear, which leads to change of shape or distortion. Differences in rock or load pressure and fluid pressure may lead to different metamorphic reactions in chemically similar systems. Whether the rock is permeable and allows fluids to migrate to lower-pressure areas is of great importance in determining chemical equilibrium. Rock or load pressure may be higher than fluid pressure but the reverse is not possible. The action of chemically active fluids is a most important factor in metamorphism, because even when they do not add or subtract material from the rocks they promote reaction by solution and redeposition. When they add or subtract material the process is called metasomatism. Probably some degree of metasomatism accompanies most metamorphism but it appears that only H_2O and CO_2 move to any significant degree so that most metamorphic reactions take place in nearly isochemical systems. This allows us to determine the parent rocks of metamorphic rocks with reasonable ease. Water is the principal chemically active fluid, and it is aided by carbon dioxide, boric acid, hydrofluoric and hydrochloric acids, and other substances.

THE CHEMICAL COMPOSITION OF METAMORPHIC ROCKS

The bulk chemical composition of metamorphic rocks is exceedingly variable. It may correspond to that of any of the igneous and sedimentary rocks, and metasomatism may produce a composition different again from these. It is, however, often possible to determine the nature of the original rock from its chemical composition even after it has been totally recrystallized and the original structure completely destroyed. We have seen how sedimentary rocks can be more extreme in their composition than igneous rocks, and this feature aids in the elucidation of their metamorphosed equivalents. Some of the chemical criteria that may be used to establish an originally sedimentary origin of metamorphic rocks are (a) an excess of alumina, which will appear as C when the norm is calculated (if C exceeds 5%, a sedimentary origin may be suspected; if C is greater than 10%, a sedimentary origin is almost certain); (b) $K_2O > Na_2O$ combined with $MgO > CaO$ (this feature is characteristic of argillaceous rocks, especially those containing appreciable amounts of illite and montmorillonite); and (c) a very high SiO_2 content (say greater than 80%, or more than 50% Q in the norm), which suggests that the rock was originally a sandstone or a chert.

The bulk chemical composition of a rock may remain constant during metamorphism (isochemical metamorphism), or it may change through the introduction and/or removal of material (allochemical metamorphism). Even isochemical metamorphism involves some transference of material, although it may be over very small distances. For example, metacrysts of manganese garnet are not uncommon in metamorphosed argillacous rocks, even though the average manganese content of the parent rock may be only 0.1%. Manganese has been transferred from the surrounding rock and concentrated in the garnet metacrysts, but the distance of migration is quite small, of the order of a millimeter. Most metamorphic migration, except for volatile components, takes place in the millimeter or centimeter range. The main impelling force in metamorphism is the addition of heat to provide the energy of activation for new mineral phases to form, small mineral grains to grow to larger size, or polymorphic transitions to take place. The distribution of elements in these reactions generally follows the rules already discussed for magmatic systems.

Allochemical metamorphism or metasomatism raises the question as to how the introduction or removal of material has taken place. Most metamorphic rocks showing extensive chemical change have limited distribution, such as contact metamorphic aureoles around intrusive magmatic bodies. Three means of material transport may be suggested: transportation in a gas phase; transportation by liquids; and transportation involving neither of these. The last process may be visualized as a migration of atoms or ions along crystal boundaries or even through solids, and its significance in metamorphism has been a subject of acute controversy. That such a process is possible is universally accepted. The controversy, like so many others in geology, rages around the magnitude of the effects this form of material transfer has actually produced in the rocks. Some workers ascribe to it the alteration of enormous volumes of material; specifically, it has been stated that granite masses of batholithic dimensions are the result of allochemical metamorphism of preexisting rocks without the intervention of liquid or gas phases. Other workers state that migration of ions in and through solids is an insignificant factor in metamorphism at best, being completely overshadowed by other types of transportation, and that in any case the requirements for extensive migration of this kind are unlikely to be realized under the usual conditions of metamorphism.

On the whole, the evidence favors the latter view. Measurements of diffusion in silicates indicate that the rate of migration of ions in solids of this kind is much too slow to produce extensive changes, even in the time available during cycles of metamorphism. Metamorphic reactions could take place in the absence of a fluid phase, but the presence of water will greatly increase the speed of such reactions and lower the minimum temperature required. A laboratory example is the formation of forsterite from MgO and quartz; when the dry solids were heated together at 1300° only about 10% reacted in two hours, whereas at 600°, in the presence of water vapor under pressure, reaction was virtually complete in the same time. Water and other volatile substances are practically omnipresent, at least in small amounts, in all rocks, and are also set free in large amounts both by igneous activity and by progressive metamorphism, thereby providing a universal and effective medium for the transport of material. Field observations and laboratory experiments indicate that metasoma-

tism is essentially the result of introduction or removal of material in a fluid phase. Korzhinsky (1955) has suggested the following sequence of relative mobility ranging from the very mobile H_2O to the essentially immobile Al_2O_3 and TiO_2:

H_2O, CO_2, S, Cl, K_2O, Na_2O, F, CaO, O_2, Fe, MgO, SiO_2, P_2O_5, Al_2O_3, TiO_2
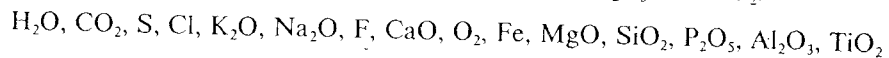
The limited diffusion of O_2 in rocks is illustrated by the fact that minerals in close association in some metamorphic rocks indicate significant differences in oxygen fugacity when they were formed. On the other hand oxygen isotope studies indicate that H_2O can migrate easily in metamorphic rocks and H_2O oxygen exchanges with rock oxygen.

Data on minor and trace elements in metamorphic rocks is limited and generalizations are few. In areas of broad regional metamorphism the distribution of trace elements, like major elements, is controlled by the nature of the parent rocks. Contact metamorphism with replacement of sediments, especially carbonates, may allow economically significant amounts of some elements to be concentrated. Elements commonly concentrated by contact metamorphism include Fe, Zn, Pb, and Cu as sulfides and W, Sn, and Ti as oxidic compounds.

THE MINERALOGY OF METAMORPHIC ROCKS

Since the chemical composition of metamorphic rocks is so varied, it is to be expected that their mineralogy will be correspondingly diverse. In addition, metamorphic rocks are formed under a wide range of temperatures and pressures, and even if no change in bulk composition takes place a mineral assemblage stable under certain (P, T) conditions may be replaced by a totally different assemblage under other (P, T) conditions. In these circumstances it is noteworthy that the mineralogy of metamorphic rocks is not more complex; this is largely due to the stability of some common minerals over considerable ranges of bulk composition and of physical conditions.

A list of some common silicate minerals of metamorphic rocks is given in Table 11.1, along with information regarding their relative compositions. The figures in Table 11.1 are derived from the ideal formulas of the minerals, all reduced to a common basis of 24 (O, OH) ions. The figures do not take into account possible ionic substitution, such as magnesium for ferrous iron or ferric iron for aluminum. The table provides a useful means for determining what additions or subtractions are necessary in the transformation of one mineral to another of related composition and for indicating the minerals that may appear in rocks of specific bulk composition.

The significant features of the important silicates with respect to metamorphism can best be discussed with regard to their structural type. Of the common tektosilicates, quartz is present in silica-rich rocks over practically the whole range of metamorphic conditions (at high temperatures and low pressures tridymite and cristobalite may form). The feldspars are common and abundant, but individual species show marked differences in occurrence. Albite is found over a wide range of conditions. Except for the occurrence of paragonite in a few schists, jadeite in eclogite, and glaucophane in glaucophane schists, albite is the principal carrier of sodium in

Table 11.1   The Composition of Minerals of Metamorphic Rocks in Atomic Proportions, Based on a Common Content of 24(O.OH) Ions

Mineral	Formula	Si	Al	Mg	Ca	Fe	Na	K
Quartz	SiO_2	12						
Andalusite, sillimanite, kyanite	Al_2SiO_5	~5	~10					
Cordierite	$Mg_2Al_4Si_5O_{18}$	~7	~5	~3				
Pyrope	$Mg_3Al_2(SiO_4)_3$	6	4	6				
Chlorite	$Mg_5Al(AlSi_3O_{10})(OH)_8$	4	~3	~7				
Enstatite	$MgSiO_3$	8		8				
Anthophyllite	$Mg_7(Si_4O_{11})_2(OH)_2$	8		7				
Talc	$Mg_3Si_4O_{10}(OH)_2$	8		6				
Serpentine	$Mg_3Si_2O_5(OH)_4$	5		8				
Forsterite	Mg_2SiO_4	6		12				
Staurolite	$Fe_2Al_9Si_4O_{23}(OH)$	4	9		2			
Chloritoid	$Fe_2Al_2(Al_2Si_2O_{10})(OH)_4$	~4	~8		~4			
Almandine	$Fe_3Al_2(SiO_4)_3$	6	4		6			
Cummingtonite	$Fe_7(Si_4O_{11})_2(OH)_2$	8			7	8		
Wollastonite	$CaSiO_3$	8				8		
Grossular	$Ca_3Al_2(SiO_4)_3$	6	4			6		
Zoisite	$Ca_2Al_3(SiO_4)_3(OH)$	~6	~6			~4		
Anorthite	$CaAl_2Si_2O_8$	6	6			3		
Diopside	$CaMgSi_2O_6$	8		4		4		
Tremolite	$Ca_2Mg_5(Si_4O_{11})_2(OH)_2$	8		5		2		
Jadeite	$NaAlSi_2O_6$	8	4				4	
Glaucophane	$Na_2Mg_3Al_2(Si_4O_{11})_2(OH)_2$	8	2	3			2	
Albite	$NaAlSi_3O_8$	9	3				3	
Potash feldspar	$KAlSi_3O_8$	9	3					3
Muscovite	$KAl_2(AlSi_3O_{10})(OH)_2$	6	6					2
Phlogopite	$KMg_3(AlSi_3O_{10})(OH)_2$	6	2	6				2

metamorphic rocks. Potash feldspar occurs more usually as microcline than as orthoclase, conditions of metamorphism being generally favorable to the crystallization of the ordered form. Anorthite is rare in metamorphic rocks, the common calcium aluminum silicate being zoisite or epidote. Plagioclase composition is often a sensitive indicator of metamorphic grade; in the lowest grade rocks it is pure albite, and the calcium content increases as the grade increases.

    Minerals of the inosilicate group, that is, the pyroxenes and amphiboles, are common and abundant constituents in metamorphic rocks. Some amphiboles and pyroxenes are virtually confined to these rocks, for example, jadeite among the pyroxenes, and anthophyllite, cummingtonite, tremolite, actinolite, and glaucophane among the amphiboles. In general, the amphiboles are typical of metamorphic rocks formed at low and moderate temperatures, the pyroxenes of those formed at higher temperatures. However, diopsidic pyroxene is often associated with hornblende or with calcite in calcareous rocks at fairly low grade.

    Minerals with layer lattices, that is, the phyllosilicates, are especially characteristic of metamorphic rocks. Talc, serpentine, the chlorites and chloritoids, and mus-

covite and biotite are common and widespread, some of them being practically confined to these rocks. The clay minerals of sedimentary rocks are particularly susceptible to recrystallization, and argillaceous rocks are thus especially suitable for tracing the initial stages of metamorphism. In this connection Yoder and Eugster (1955) have recognized a series of subtle changes in the crystal structure of the clay mica, from a one-layer, randomly stacked polymorph to a one-layer ordered polymorph to a two-layer ordered polymorph during progressive metamorphism. Similar changes can be expected in other phyllosilicates, such as the chlorites.

The nesosilicates are common and abundant in metamorphic rocks, and some of them, such as the garnets, epidote, and the aluminum silicates, are especially typical of such rocks. The nesosilicates in general are closely packed structures and would be expected to show marked stability under high pressure. The aluminum silicates—kyanite, sillimanite, and andalusite—are found in metamorphic rocks with high aluminum content. Stability relations of these three polymorphs are illustrated in Figure 11.1. Kyanite has a considerably higher density than the other polymorphs and indicates conditions of high pressure, although it occasionally occurs as a vein mineral. Andalusite is characteristic of thermally metamorphosed rocks and appears to be unstable under stress. Sillimanite is widespread in high-grade metamorphic rocks. Under extremely high temperatures typified by the contact action of basic lavas the mineral mullite, which resembles sillimanite closely but is of somewhat different composition, may appear. The garnets occur frequently in metamorphic rocks. Their composition is, of course, conditioned by the bulk composition of the rock, but they are also sensitive indicators of metamorphic grade. In the lowest grades of meta-
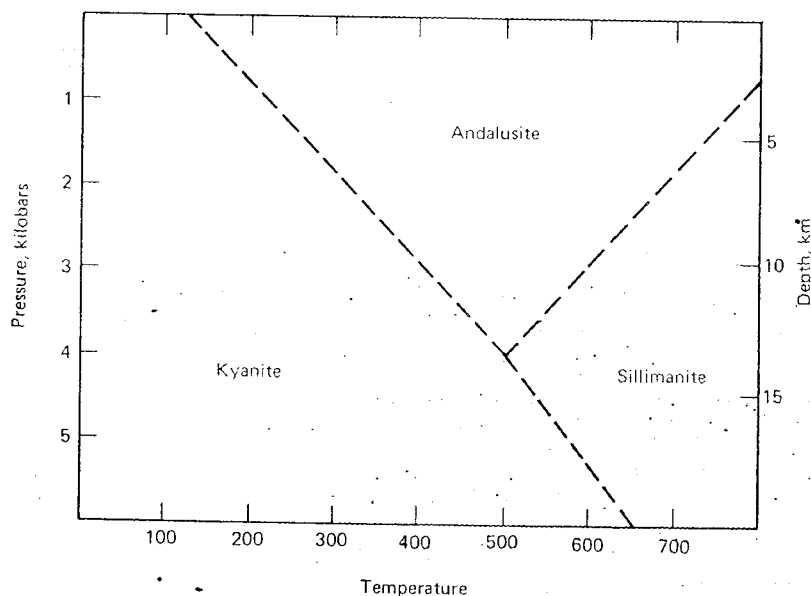
Figure 11.1 Stability fields of Al_2SiO_5 minerals (phase boundaries approximately located).

morphism the only garnet found is the manganese variety spessartine. At somewhat higher grades of regional metamorphism almandine may be formed, whereas garnets rich in the pyrope component are formed only at high metamorphic grades. The calcium-iron garnet andradite is characteristic of limestones that have been metasomatically altered by iron-rich emanations. The zoisite-epidote group is a common constituent of metamorphic rocks of low and medium grade. Minerals of this group are important carriers of calcium and aluminum in such rocks.

In general it can be said that the distinctive silicates of metamorphic rocks are the inosilicates and the phyllosilicates. This is not due to chance; they are minerals with fairly high density, the formation of which is favored by increased pressure; their structures generally tolerate considerable atomic substitution and may therefore form under variable bulk composition of the rocks. Some nesosilicates, especially minerals of the garnet and epidote groups, are also characteristic of many metamorphic rocks. Many of the tektosilicates or framework structures, on the other hand, show a marked instability under conditions of metamorphism, probably because of their distinctly open lattices.

The role of aluminum in the silicate minerals of metamorphic rocks deserves mention. We have seen that aluminum is capable of either sixfold or fourfold coordination with oxygen; in sixfold coordination it is structurally similar to magnesium or ferrous and ferric iron, whereas in fourfold coordination it acts like silicon. A distinct correlation exists between the type of coordination of aluminum and the grade of metamorphism. In general, four-coordination of aluminum is promoted by increasing temperature; thus the amount of aluminum replacing silicon in such minerals as the amphiboles, pyroxenes, and micas increases at higher metamorphic grades. The effect of pressure is less obvious, but sixfold coordination is more economical of space and should be favored by high pressures. Minerals containing four-coordinated aluminum are characteristic of igneous rocks and the products of thermal and high-grade regional metamorphism, whereas in sedimentary rocks and low- to medium-grade metamorphic rocks aluminum typically occurs in six-coordination.

Many other minerals besides the silicates discussed here occur in metamorphic rocks, but generally in quite minor amounts. An exception should be made for calcite and dolomite, which make up the greater part of some metamorphic rocks. Their stability is a direct function of the partial pressure of carbon dioxide during metamorphism, and they may thus be stable over a wide range of temperatures and pressures, provided sufficient carbon dioxide is present to prevent decomposition. Magnesium carbonate is much more readily decomposed than calcium carbonate, and therefore calcite is stable to higher metamorphic grades than dolomite.

The calculation of the quantitative mineralogical composition of a metamorphic rock from its chemical analysis is a much more complex procedure than for an igneous or sedimentary rock. Metamorphic rocks may crystallize over a wide variety of physicochemical conditions, and the same chemical composition may correspond to several different mineralogical compositions, depending upon the specific conditions (cf. Table 11.2). Of course, the problem is greatly simplified if the qualitative mineral composition is known from microscopic examination. Barth has developed procedures for calculating normative mineralogical compositions of metamorphic

rocks, similar in principle to the norms of igneous rocks; he distinguishes catanorms, mesonorms, and epinorms for rocks of high-, medium-, and low-grade metamorphism respectively.

THE STABILITY OF MINERALS

Stability is an environmental property. For example, calcite is said to be stable at ordinary temperatures and pressures. So it is, but only in the sense that under specific conditions it remains unchanged indefinitely. If it comes in contact with dilute hydrochloric acid, it becomes far from stable. Stability must therefore be defined not only in respect to pressure and temperature, but also as regards chemical environment. Strictly speaking one cannot speak of the stability of a single phase (mineral) in a rock, independent of the associated minerals and the pore fluids. Most discussions of particular minerals or mineral associations assume an indifferent chemical environment, in which stability is a function of physical conditions only.

Three states may be distinguished: stability, metastability, and instability. A stable mineral association is that which has the lowest possible free energy under the particular circumstances and which therefore has no tendency to change. A mineral association is unstable when it is not the association with the lowest free energy under the specific conditions, and when the rate of change to an association with lower free energy is appreciable. A metastable association is one with more than the minimum free energy for the system, but in which the rate of change to an association with lower free energy is so slow as to be undetectable—an input of energy is required to make the transformation take place at a finite rate. Stability therefore involves two independent factors, thermodynamics and kinetics.

THERMODYNAMICS OF METAMORPHISM

The thermodynamics of metamorphism can be expressed in terms of the transformation of a mineral association $A + B + C \cdots$ into a different association $L + M + N \cdots$ according to the equation

$$A + B + C \cdots = L + M + N \cdots$$
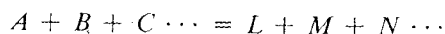
For this change to take place the net free energy of $L + M + N \cdots$ must be less than that of $A + B + C \cdots$; that is, the free energy change must be negative. This condition can be expressed in the following way:

$$\Delta G = (G_L + G_M + G_N \cdots) - (G_A + G_B + G_C \cdots) < 0$$

However, in general the free energy change is obtained indirectly, usually by the application of the second law of thermodynamics, as expressed in the following form

$$\Delta G = \Delta H - T \Delta S$$

in which ΔH is the enthalpy or heat of reaction and ΔS the entropy change.

The application of the simple thermodynamic equation to the consideration of actual reactions is complicated by the variation of ΔS and ΔH with temperature and

pressure, and in order to study a reaction over a range of temperatures and pressures one must have data on compressibility, specific heat, and thermal expansion for the different phases. In addition, ΔS and ΔH have not been measured for many reactions of geological significance. Such measurements would provide basic data for significant advances toward the elucidation of fundamental geochemical problems. However, qualitative statements based on Le Châtelier's principle are often possible. Increased pressure will cause a reaction to run in a direction such that the total volume of the system decreases. Increasing temperature favors endothermic reactions. Since the volume change accompanying metamorphic reactions (except those involving gas phases) is generally small in comparison to the heat of reaction, temperature changes are more significant than pressure changes in displacing equilibrium.

Metamorphic reactions may be divided into four general types.

1. Solid-solid reactions involving no liberation or addition of volatiles.
2. Dehydration-hydration reactions involving the loss of H_2O with a temperature increase or the hydration of rocks by the addition of heated water.
3. Decarbonation reactions involving the loss of CO_2.
4. Oxidation-reduction reactions often brought about by the loss of oxygen by the production of volatile CO or CO_2 from reduced carbon species.

A typical example of solid-solid reactions is illustrated by the polymorphs of Al_2SiO_5 shown in Figure 11.1. Many studies of this system confirm the general shape of the phase diagram, but the location of the triple point is not unambiguously known because of the sluggishness of the reaction rates.

The slopes of equilibrium curves can be predicted from the Clapeyron relation

$$\frac{dP}{dT} = \frac{\Delta S}{\Delta V}$$

Thermodynamic studies have been applied to the investigation of the formation of jadeite ($NaAlSi_2O_6$) in metamorphic rocks. Jadeite is a rare mineral and in spite of its simple chemical composition its laboratory synthesis proved difficult. It is, of course, a phase in the system Na_2O—Al_2O_3—SiO_2 and is intermediate in composition between albite and nepheline. The thermodynamic data show that jadeite is a stable phase in this system at 25° and 1 atm pressure. That it is a rare mineral is evidently not because of any lack of stability but probably because of the extreme sluggishness at low temperatures of the reactions that will form it. The correctness of these deductions from the thermodynamic data has been confirmed by the discovery of jadeite as an important constituent of Californian graywackes, which have evidently been metamorphosed at quite low temperatures. In these rocks jadeite has apparently been formed by the reaction albite = jadeite + quartz. For this reaction at 25° and 1 atm pressure, ΔG has a small positive value which diminishes with increasing pressure and becomes negative above about 1600 atm; that is, jadeite and quartz would then tend to form from albite. Such a pressure corresponds to a depth of about 7,000 meters, and the jadeite-bearing graywackes were probably metamorphosed at this depth or greater.

A well-studied metamorphic system in which a volatile phase is a component of the reaction is the formation of wollastonite from calcium carbonate and silica according-to-the equation

$$CaCO_3 + SiO_2 = CaSiO_3 + CO_2$$
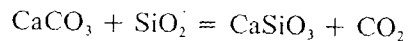
The thermodynamics of this reaction and its significance for the metamorphism of siliceous limestone was first discussed by Goldschmidt, and he derived therefrom information regarding pressure-temperature conditions during metamorphism. With the aid of modern high temperature-high pressure equipment Harker and Tuttle have succeeded in determining the equilibrium curve for the reaction experimentally (Figure 11.2). The general features deduced by Goldschmidt are confirmed, although the reaction takes place at a lower temperature than he predicted. At zero pressure of carbon dioxide the formation of wollastonite may take place at temperatures below 400°. As may be predicted from the law of mass action, this temperature increases with increasing pressure, reaching 750° at 2000 atm. Figure 11.2 shows that for conditions above curve AB the formation of wollastonite from calcium carbonate and silica can take place spontaneously ($\Delta G < 0$); below curve AB the reaction is reversed, and calcium carbonate and silica can be formed from wollastonite and carbon dioxide; curve AB itself represents equilibrium conditions ($\Delta G = 0$).

Goldschmidt's computations and Harker and Tuttle's experimental work were performed on a closed system in which the pressure was that of the CO_2. Barth has pointed out that under geological conditions this situation is improbable, since during the reaction of calcium carbonate and silica the carbon dioxide produced will tend to diffuse away. Wollastonite is denser than either calcite or quartz, and so in terms of the solids the reaction proceeds with a decrease in volume, and hence if the CO_2
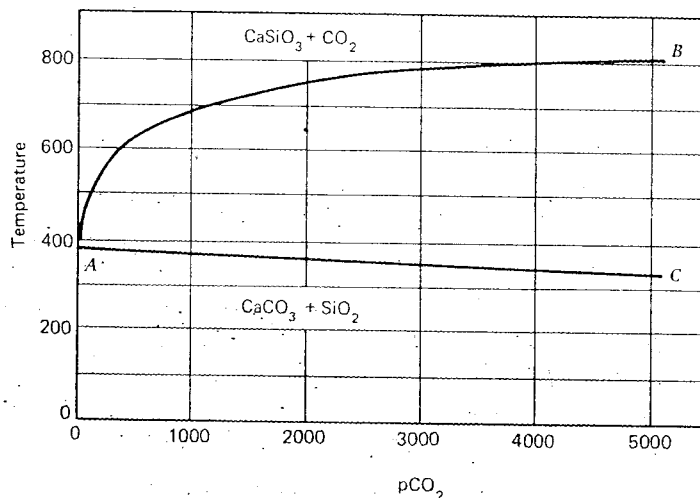
Figure 11.2   Pressure-temperature curves for the reaction $CaCO_3 + SiO_2 = CaSiO_3 + CO_2$.

escapes the temperature of formation of wollastonite will decrease with increasing pressure. Curve AC is the quantitative expression of this decrease, as calculated by Barth by means of the Clapeyron equation. This example illustrates the importance of distinguishing between open and closed systems for reactions involving volatile components. Such reactions are common in metamorphism.

Many metamorphic reactions are dehydration reactions. Their thermodynamic characteristics are similar to the decarbonation reaction discussed above. Any natural metamorphic reaction involving the loss of the volatiles H_2O and/or CO_2 has a number of composition–pressure possibilities. The mobility of H_2O and CO_2 can be different in different cases. In an open system where fluids can escape the fluid pressure will be less than the rock pressure; such a system will fall between curves like AC and AB in Figure 11.2. If the mobility of H_2O is not great with respect to the dehydration reactions and the H_2O phase is retained between mineral grains, the H_2O pressure may equal the rock pressure. If pelitic or clay-rich sediments occur with calcareous ones, metamorphic reactions may release a fluid containing both H_2O and CO_2. Under these circumstances the fluid pressure may either be equal to or less than the rock pressure but the H_2O pressure in either case will be less than the total fluid pressure.

In addition any metamorphic fluid will contain dissolved salts and therefore the chemical activity of either H_2O or CO_2 will be difficult to predict with a great degree of precision. In order to construct phase diagrams of geologic interest assumptions are usually made to fix selected variables and the variation of the other components under these conditions are illustrated. Figure 11.3 shows two equilibrium curves at constant total pressure for the wollastonite reaction with variable partial pressures of CO_2. Graphs of this type are isobaric plots, and may be used to illustrate reactions in mixed H_2O-CO_2 systems, which more likely approximate actual systems than do simple decarbonation studies.

In addition to the addition or loss of H_2O and CO_2, metamorphic rocks often show changes in the ferrous-ferric ratios. This ratio tends to decrease with increasing temperature of metamorphism and hence indicates a change of the partial pressure of $O_2 : PO_2$. Indicators of PO_2 in metamorphic rocks include the oxides of iron. Reactions of the type

$$6\ Fe_2O_3\ \text{(hematite)} = 4\ Fe_3O_4\ \text{(magnetite)} + O_2$$
$$2\ Fe_3O_4\ \text{(magnetite)} = 6\ FeO\ \text{(wüstite)} \quad\ \ + O_2$$
$$2\ FeO\ \text{(wüstite)} \quad = 2\ Fe\ \text{(iron)} \qquad\quad + O_2$$
$$Fe_3O_4\ \text{(magnetite)} = 3\ Fe\ \text{(iron)} \qquad\quad + 2O_2$$
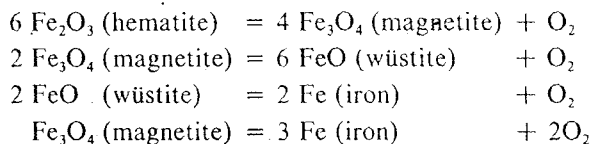
act as natural buffers for PO_2. Hematite and magnetite are common constituents of metamorphic rocks but wüstite and native iron are not. Figure 11.4 illustrates the PO_2 temperatures relationships for the above reactions. It may be inferred that the PO_2 in common metamorphic rocks is in the range of 10^{-10} to 10^{-35} bar. Also included in Figure 11.4 are dissociation curves for H_2O and CO_2 under different conditions. If moderate quantities of H_2O are present, the PO_2 will be high enough for hematite to be stable. The common presence of magnetite indicates lower H_2O
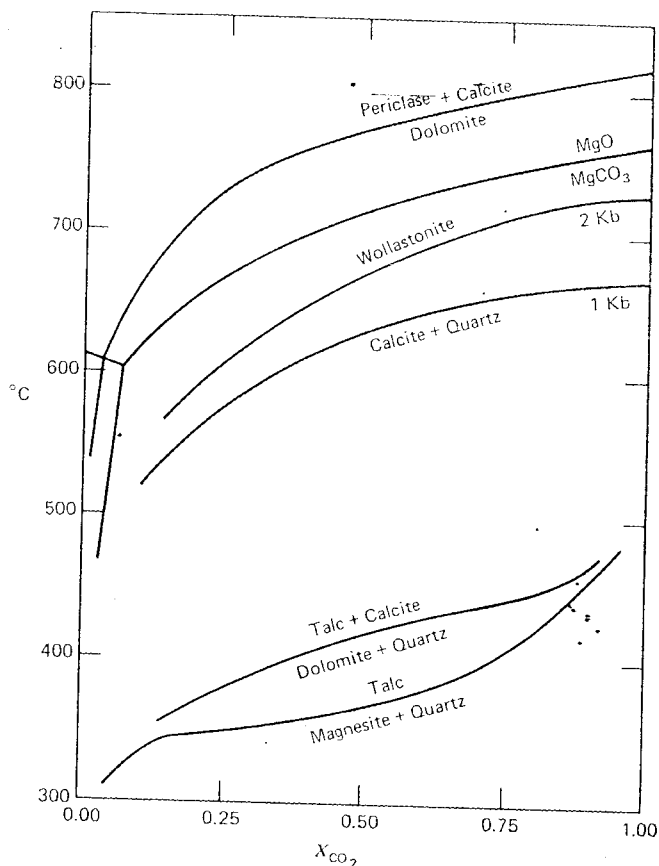
Figure 11.3 Isobaric equilibrium curves of various metamorphic reactions. X_{CO_2} is the mole fraction of CO_2 in the fluid phase which consists of $CO_2 + H_2O$. The fluid pressure P_f is 1 kb, except where indicated otherwise. (Winkler, 1979, p. 126.)

contents. The presence of carbonaceous matter in metamorphic rocks is of importance in buffering PO_2 according to the reaction .

$$C + O_2 = CO_2$$
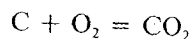
The production and migration of CO_2 is a major control of PO_2 over a great range of metamorphic reactions and explains the common occurrence of magnetite. Secondary reactions of H_2O and CO_2 also account for concentrations of CO and CH_4 and provide for a PH_2 buffer system as well as PO_2.

Thermodynamic data thus enable predictions concerning the relative stability of possible phases under specified conditions of temperature, pressure, and chemical environment. Knowledge of the free-energy change accompanying any reaction is adequate for determining whether the reaction can take place. Those reactions take
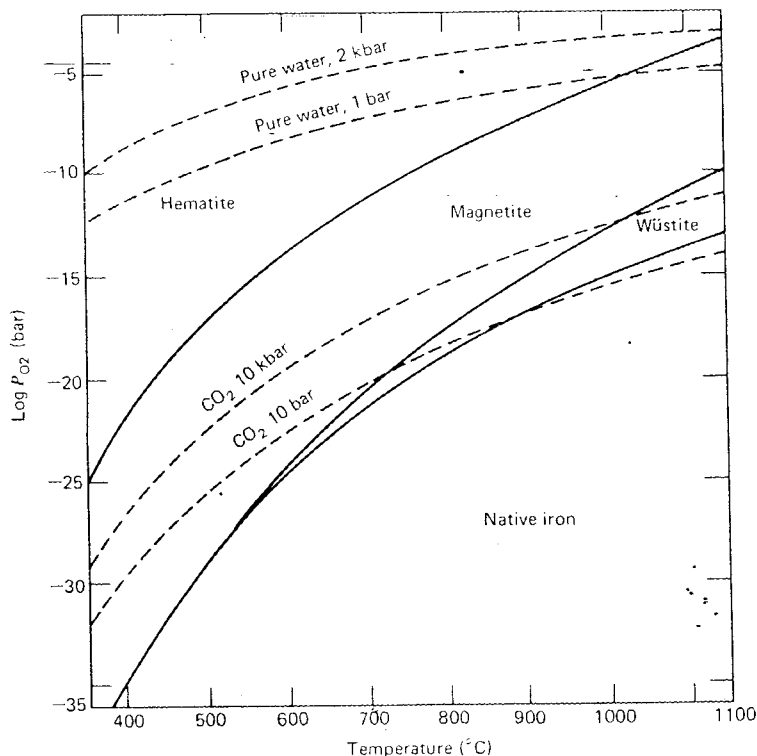
Figure 11.4   Stability relations of oxides of iron and native iron. The
effect of pressure on solid phases is neglected.(Miyashiro,
1973, p. 52.)

place that lead to a decrease in the free energy of the system as a whole. Thus during
metamorphism the tendency is toward the transformation of the rock material into
a mineral association that for the specific conditions has the lowest possible free
energy. However, the rate at which such transformation will proceed cannot be
determined from the thermodynamic data but is dependent on other factors.

KINETICS OF METAMORPHISM

The significance of kinetics in metamorphism can be illustrated by the simple case
of recrystallization of a limestone. A fine-grained aggregate of calcite can alter spon-
taneously to a coarse-grained aggregate, since the change will result in a reduction
of the net surface energy. Yet some fine-grained limestones have remained essentially
unchanged for millions of years, often in spite of deep burial and the accompanying
temperature increase. Something more than a potential decrease in free-energy con-
tent is evidently required to cause a metamorphic change to take place.

The kinetics of heterogeneous reactions are not well understood, but qualitative statements can be made with some degree of assurance. The ionic groupings found in minerals are all relatively stable entities, and before they can be rearranged to give different minerals they must undergo some kind of dislocation. Energy is required for this, and only when energy is available in appreciable amounts can the rate of reaction become significant. The factor is an exponential function of E/RT, in which E is known as the *activation energy*. This function makes the activation energy the most important single factor determining reaction velocities and shows that they are increased by a certain amount per unit temperature rise. The usual statement that a reaction rate is doubled by a rise in temperature of $10°$ is a rough generalization and implies a uniform energy of activation of about 13,000 cal/mole; measured energies of activation show a considerable range but are of this order of magnitude.

Temperature is thus the most significant factor in determining the rate of a reaction. However, any agent that weakens the bonding within the phases present will give the reacting units greater freedom of movement and thereby decrease activation energies and accelerate reaction. The influence of water, a powerful solvent and ionizing agent, is particularly significant in mineral transformations, which are essentially ionic in character. Water or water-rich fluids may be essential for the catalysis of metamorphic reactions.

Although rise in temperature is perhaps the most important source of activation energy, it is by no means the only one. Shear is an important source of energy in metamorphism and is probably responsible for much reconstitution that would not take place in its absence. This is borne out by field observations that suggest the formation of low-grade schists has taken place at temperatures that may not exceed those existing in deep geosynclines, where rocks are indurated but not recrystallized. The contrast is evidently a reflection of the differing effects of uniform load and directed pressure.

The activation energy of the reactions taking place during metamorphism may be expected to vary according to the nature of the reaction and to be considerably influenced by the type of structural changes involved. Studies of polymorphic changes have shown that the rate of such transformations depends greatly on the degree of rearrangement of the structural units. A typical example is the rapid inversion of low-quartz to high-quartz compared with the sluggish transformation of high-quartz to tridymite. By analogy it may be expected that the change from chlorite to biotite, for example, probably has a much lower activation energy than the change from chlorite to garnet; in the first change large segments of the chlorite lattice can be directly incorporated in the biotite lattice, whereas in the latter change the sheets of silicon-oxygen tetrahedra must be completely broken down before the garnet structure can be built up from them.

In addition to the above factors, the rate of reconstitution of a rock will also be influenced by the physical condition of the reacting material, particularly the size of the individual grains and the intimacy of their mixture. The reacting units come together more easily the smaller the grain size and the more complete the mixing; that is, fine-grained rocks will respond to metamorphism more rapidly than those of

coarser grain. Here stress has an important mechanical effect, by grinding the constituent minerals together and thereby crushing them and bringing them into closer contact.

Even though metamorphic reactions may proceed slowly we must not forget that long periods of time are available. Calculations on the thermal regimes surrounding plutonic intrusions have shown that high temperatures are retained around the intrusion for tens of millions of years. Studies of argon retention in metamorphic minerals indicate that temperatures exceeding those required to seal against argon diffusion may exist over 100 million years in some regions of regional metamorphism. Even so, the rate-controlling step in metamorphism appears to be related to material transport rather than temperature fluctuation. Studies of solid-state diffusion coefficients in silicates show that even in the range of 600° to 1000°C diffusion coefficients have values in the range 10^{-10} to 10^{-16} m^2/sec.

If mean diffusion distances $\bar{x}$ are estimated using the relationships $\bar{x}^2 = 2\,Dt$, where D is the diffusion coefficient and t is time, even over periods of 100 million years atomic migration is in the range of centimeters to meters. For elements to move over greater distances intergranular or fluid-assisted diffusion must take place. If a fluid phase is removed metamorphic reactions virtually cease. For this reason high-temperature metamorphic systems are preserved on cooling due to the difficulties of reintroducing water to anhydrous systems.

The importance of reaction rates can be illustrated by a consideration of the kinetics of the wollastonite reaction discussed in the preceding section. The thermodynamic data show that wollastonite can form from calcium carbonate and silica at temperatures as low as 380°. However, in the laboratory the reaction between calcium carbonate and silica is generally not detectable below about 500°. On the basis of the thermodynamic data alone we might conclude that the absence of wollastonite in a limestone of suitable composition indicates that its temperature had never exceeded 380°. This conclusion would not be justified, however, since the rate of formation of wollastonite may not become finite, even under favorable circumstances, until higher temperatures are reached.

The sluggishness of many reactions at low to moderate temperatures undoubtedly results in the persistence of some minerals under conditions in which they should be changed to others with a lower free energy. This is particularly true in rocks that have been metamorphosed at high temperatures. Their mineralogy is more or less "frozen" at these high temperatures and often shows little if any effect of retrogressive metamorphism during cooling. In general, response in rocks to rising temperatures will be more rapid than response to falling temperatures because of the accelerating effect of temperature increase on reaction rates.

METAMORPHISM AND THE PHASE RULE

The most fruitful line of approach to the geochemistry of metamorphism has been through the principles of equilibria in heterogeneous systems as codified by the phase rule. The study of metamorphic rocks indicates that equilibrium is generally attained during metamorphism, although the evidence for it is usually indirect. In igneous

rocks it is often possible to decide such a question by direct reference to the results of laboratory experimentation. Experimental observation of equilibria in silicate systems at the temperatures and pressures of metamorphism has not yet progressed to the same extent. In general, the evidence for equilibrium in the phase assemblages of metamorphic rocks rests on such criteria as their simple mineralogical composition and the tendency for certain typical mineral associations to recur in rocks of the same general composition, despite widely different age and locality. By the phase rule the number of minerals capable of existing together as phases of a system at equilibrium is limited by the number of components in the rock. We find that usually even fewer phases than the maximum number allowed by the phase rule are present. Many chemically complex metamorphic rocks, such as amphibolite (hornblende-plagioclase) and mica schist (muscovite-chlorite-albite-quartz), are extremely simple mineralogically. This is evidently because the individual elements that appear in the composition of the rock are not distinct components in the thermodynamic sense; that is, ferrous iron and magnesium often act as a single component and, to a lesser degree, aluminum and ferric iron, and sodium and calcium.

As mentioned above, the approach that has proved so fruitful in the elucidation of the evolution of the igneous rocks—the laboratory investigation of phase relations in silicate systems—has also been extended to conditions corresponding to those obtaining during metamorphism. The first extensive research of this kind to appear is that for the system MgO—SiO_2—H_2O (Figure 11.5).

Difficulties in studying metamorphic systems mainly result from the technical problems in carrying out equilibrium studies involving volatiles, and the sluggishness
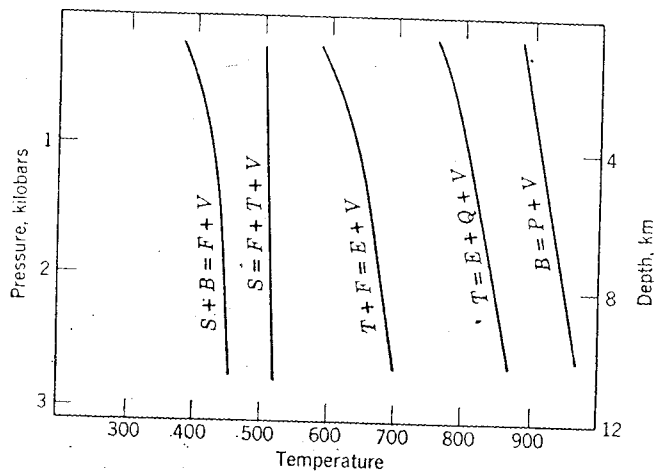
Figure 11.5 Pressure-temperature curves of univariant equilibrium in the system MgO-SiO_2-H_2O. The equation on each curve indicates the reaction to which the curve refers; B = brucite, E = enstatite, F = forsterite, P = periclase, Q = quartz, S = serpentine, T = talc, V = water vapor. (Bowen and Tuttle, *Bull. Geol. Soc. Am.* 60, 447, 1949)

of many reactions at temperatures and pressures corresponding to those of metamorphism. A useful approach has thus been from observations on metamorphic rocks. Geologists have studied the products of completed reactions, namely the mineral associations of metamorphic rocks, and much progress has been made, especially in determining compatible and incompatible minerals and mineral associations. Here again, as in the investigation of mineral formation in igneous rocks, real advances date from the introduction of physicochemical principles, especially the phase rule, as a guide in the study of reactions in rocks. The credit for applying these principles to metamorphic rocks goes in the first place to Goldschmidt, for his classic work on contact metamorphism in the Oslo region published in 1911, and secondly to Eskola, who in 1915 applied the same principles to regional metamorphism in his study of the relationship between mineralogical and chemical composition in the metamorphic rocks of the Orijärvi region. These and other studies led to the formulation of the facies principle for the classification of metamorphic rocks.

THE FACIES PRINCIPLE

Ideally, a genetic classification of metamorphic rocks would be based on the temperature and pressure at which the mineral assemblage originated. Eskola recognized that the mineral associations are indicative of conditions of formation as well as of the chemical composition of the rock itself. He exemplified this principle by selecting different rocks with similar chemical composition and showing that the mineralogical composition of each was distinctive (Table 11.2).

In the rocks of Table 11.2 the bulk composition is much the same throughout, and (assuming equilibrium has been reached in each case) the different mineral associations can only represent different conditions of crystallization. This phenomenon is the basic idea behind Eskola's facies classification. Each of the rocks in Table 11.2 represents a distinct facies. The mineral assemblage of a rock reflects the physical conditions under which it developed and is thus the criterion by which the facies may be recognized. The term is analogous to the stratigraphic facies, which comprises sediments with characters pointing to a genesis under similar circumstances. Eskola formulated the concept of mineral facies in the following words:

> A mineral facies comprises all the rocks that have originated under temperature and pressure conditions so similar that a definite chemical composition has resulted in the same set of minerals, quite regardless of their mode of crystallization, whether from magma or aqueous solution or gas and whether by direct crystallization from solution (primary crystallization) or by gradual change of earlier minerals (metamorphic crystallization).

Ramberg, who has done much work on the mineral facies concept, expresses it succinctly: "Rocks formed or recrystallized within a certain (P, T) field, limited by the stability of certain critical minerals of defined composition, belong to the same mineral facies."

These definitions show that the concept can be applied to any rock at all, sedimentary, igneous, or metamorphic, but in practice it has found its greatest utility in

Table 11.2 Mineralogy of Rocks of Gabbroic Composition from Different Facies

Diabase Facies		Pyroxene Hornfels Facies		Amphibolite Facies		Epidote-Amphibolite Facies		Greenschist Facies		Granulite Facies		Eclogite Facies		Glaucophane Schist Facies	
Diabase (East Karelia)		Hornfels (Oslo Region)		Amphibolite (Kisko, Finland)		Epidote-Amphibolite (Sulitelma, Norway)		Chlorite Schist (Sulitelma, Norway)		Norite Granulite (North Finland)		Eclogite (Burgstein, Tyrol)		Glaucophane Schist (Scalea, Italy)	
Plagioclase	48.4	Plagioclase (Mean An_{40})	48	Plagioclase (Mean An_{13})	26.5	Plagioclase (An_9)	42.8	Albite	39.9	Plagioclase	49.5	Omphacite	48.5	Glaucophane	54.3
Pigeonite	37.4	Hypersthene	17	Hornblende	71.5	Hornblende	42.2	Chlorite	29.4	Hypersthene	25.3	Garnet	50.5	Lawsonite	26.8
Hornblende, mica, etc.	14.2	Diopside	18	Quartz	2.0	Clinozoisite	12.3	Epidote	23.0	Diopside	9.6	Others	1.0	Sericite	15.6
		Biotite, iron ore, etc.	17			Others	3.4	Others	7.2	Orthoclase	7.1			Others	3.1
										Others	7.4				

the study of metamorphic rocks. A mineral facies comprises all rocks that have reached chemical equilibrium under a particular set of physical conditions. In effect, the facies concept is an expression of the phase rule stated in the following way: In any system at equilibrium the number and composition of the different phases will depend only on the bulk composition of the system and the temperature and pressure at which equilibrium was reached. Facies are recognized and defined by the occurrence of critical minerals or mineral associations that are characteristic of the facies in question and that do not appear in any other facies. Relatively few minerals are sufficiently sensitive to be critical in this sense; many—quartz, calcite, albite, and others—are stable over such a range of conditions that they may appear in several quite distinct facies.

The actual ranges of temperature and pressure characteristic of each facies are not known with a high degree of precision. Estimates by different investigators vary considerably, and indeed are frequently somewhat contradictory. Judging from the geological evidence, such as the transition from anthracite to graphite accompanying the transition from indurated sedimentary rocks to schists of the greenschist facies, the lower temperature boundary of this facies is at least 200°C. If we put the lower boundary of the epidote-amphibolite facies at the incoming of biotite, a temperature of 400° may be a reasonable estimate. The boundary between the epidote-amphibolite facies and the amphibolite facies can perhaps be placed near 500°C, and that between the amphibolite facies and the granulite facies at about 650°C. Granulite facies rocks are notably "dry" in the sense that hydroxyl-bearing minerals are essentially absent; an appreciable amount of available water would probably induce melting under conditions of this facies. The progressive metamorphism of amphibolite facies rocks may result either in partial melting and the formation of a magma (if H_2O is available), or recrystallization in the granulite facies (if H_2O is unavailable or can escape from the system).

Because different facies represent different temperature-pressure fields, their mutual relationships can be represented by a diagram using temperature and pressure as axes (Figure 11.6). If the pressure is due to rock load, it can be correlated with depth of burial, as indicated on the right of the diagram. This figure also serves to illustrate the diffuse boundary between metamorphism and diagenesis, on the one hand, and between metamorphism and magmatism on the other. The minimum melting curves of "granite" and "basalt" define the broad zone within which magmas of different composition crystallize. To the left of the melting curve of "granite" no remelting of normal silicate rocks can be expected. In the region between this curve and the melting curve of "basalt" the amount of remelting (anatexis) a rock will suffer will depend largely on its composition and the presence or absence of volatiles. Pure quartzite and alkali-poor slates may remain essentially solid at even higher temperatures, but other rocks will probably produce some melt in this region. It is clear that no sharp division between metamorphism and magmatism is possible.

In Figure 11.6 the field of metamorphism is limited on the low-temperature side by a line representing a minimum thermal gradient in the crust. This is based on a temperature of 300° at a depth of 30 km. The normal facies of regionally metamorphosed rocks are indicated as lying along a considerably steeper thermal gradient.
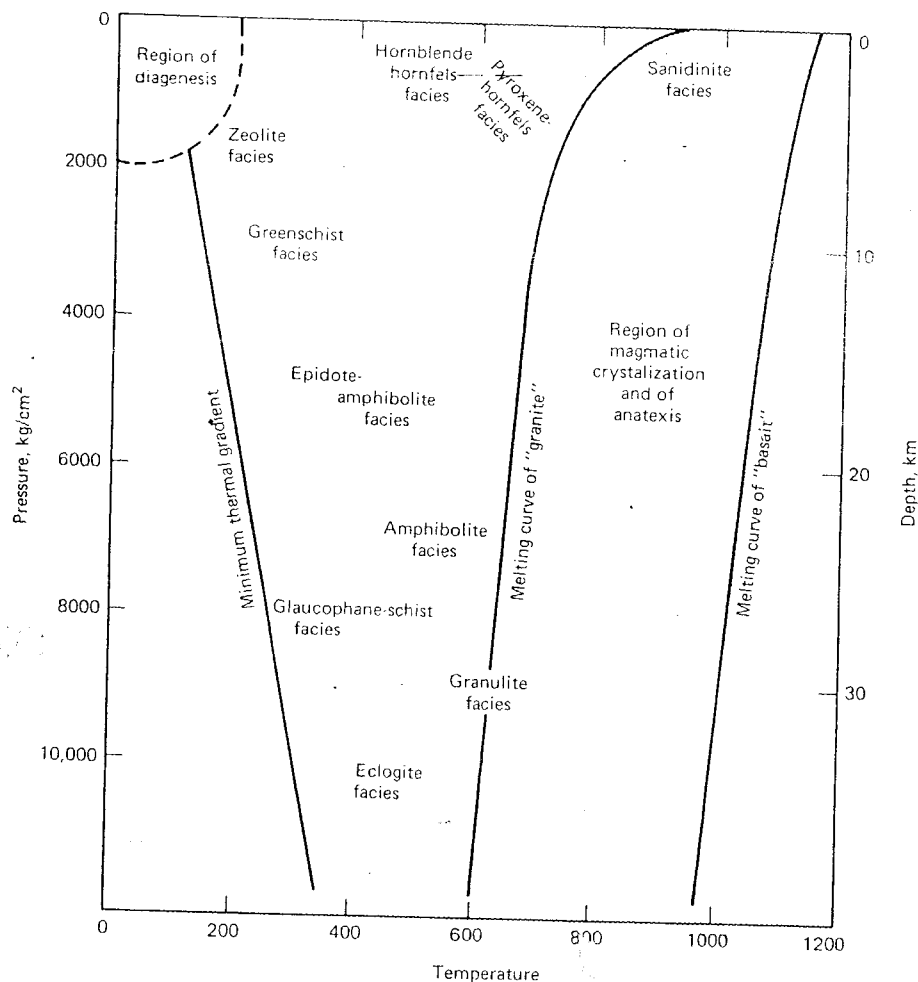
Figure 11.6 The principal metamorphic facies in relation to temperature and pressure.

The reactions accompanying progressive regional metamorphism are generally endothermic, and the process thus requires a considerable input of heat, probably more than that available under the conditions of the minimum thermal gradient. Orogeny, regional metamorphism, and igneous activity are broadly associated in space and time, and they can all be regarded as geological manifestations of an input of energy, mainly thermal.

The greenschist, epidote-amphibolite, amphibole, and granulite facies may be called the normal facies of regionally metamorphosed rocks, arranged in order of increasing grade of metamorphism. The eclogite facies and glaucophane-schist fa es are believed to originate under unusually high pressure, being characterized by minerals with high density (garnet, jadeitic pyroxenes, and lawsonite). The sanidinite facies to the right of the pyroxene-hornfels facies is exemplified by xenoliths in basic

lavas, in which the temperature has been very high, leading to incipient vitrification and the development of such minerals as mullite and sanidine.

The characteristic mineralogical features of the different facies can most clearly be illustrated by means of triangular composition diagrams, similar in principle to the diagrams used to illustrate three-component systems. In such diagrams each of the apices represents a particular component; those components chosen are mainly responsible for the observed mineralogical variation within the facies. Such diagrams were introduced by Eskola and are usually *ACF* diagrams in which *A* is Al_2O_3 + Fe_2O_3, *C* is CaO and *F* is (Fe,Mg)O. The effect of silica on the possible mineral assemblages can be illustrated by comparing two *ACF* diagrams, one (the most usual case) for rocks with excess SiO_2 (present as free quartz) and the other for rocks deficient in SiO_2. In the same way diagrams have been constructed for rocks of special compositions, viz., *AKF* diagrams in which *K* represents K_2O. *ACF* diagrams illustrating the different facies are given in Figure 11.7.

The critical minerals for the sanidinite facies are sanidine, clinoenstatite-clinohypersthene and their mix-crystals with diopside-hedenbergite (pigeonite) and (at very high temperatures) the aluminum silicate mullite. In limestones and dolomites metamorphosed under these conditions a considerable number of unusual minerals have been described, such as larnite (Ca_2SiO_4), rankinite ($Ca_3Si_2O_7$), and merwinite ($Ca_3MgSi_2O_8$).

A critical mineral combination for the pyroxene-hornfels facies (in common with the granulite facies) is the pair hypersthene-diopside, which does not form mix-crystals under the conditions of this facies. This distinguishes the pyroxene-hornfels facies from the sanidinite facies. Another difference is that (Mg,Fe)SiO_3 in this facies is always orthorhombic. Sanidine does not appear, potash and soda felspar forming distinct phases. The potash feldspar is orthoclase, not microcline, in harmony with the idea that orthoclase is a disordered form and microcline an ordered form.

The granulite facies comprises rocks that have been subjected to high-grade regional metamorphism. It is characterized by the absence of micas; instead of miscovite we find sillimanite (or kyanite) and orthoclase, instead of biotite, garnet and orthoclase. Calcite is probably the stable phase at high lime compositions instead of wollastonite, provided sufficient CO_2 is present.

In the amphibolite facies amphibole always appears, provided the bulk composition allows it. The combination hornblende-plagioclase is critical. Its high-temperature limit is marked by the appearance of diopside and hypersthene in place of hornblende, and its low-temperature limit by plagioclase composition, which changes to albite with decreasing metamorphic grade.

For the epidote-amphibolite facies the combination hornblende-albite-epidote is critical. Under the conditions of this facies the anorthite component of plagioclase is converted to zoisite or epidote; the plagioclase thus becomes high in the albite component. Hornblende is still stable.

At metamorphic grades lower than those for the epidote-amphibolite facies hornblende is unstable, and in its place one finds more chlorite and epidote, or, if the partial pressure of CO_2 is high, dolomite or magnesite plus quartz. Carbon dioxide
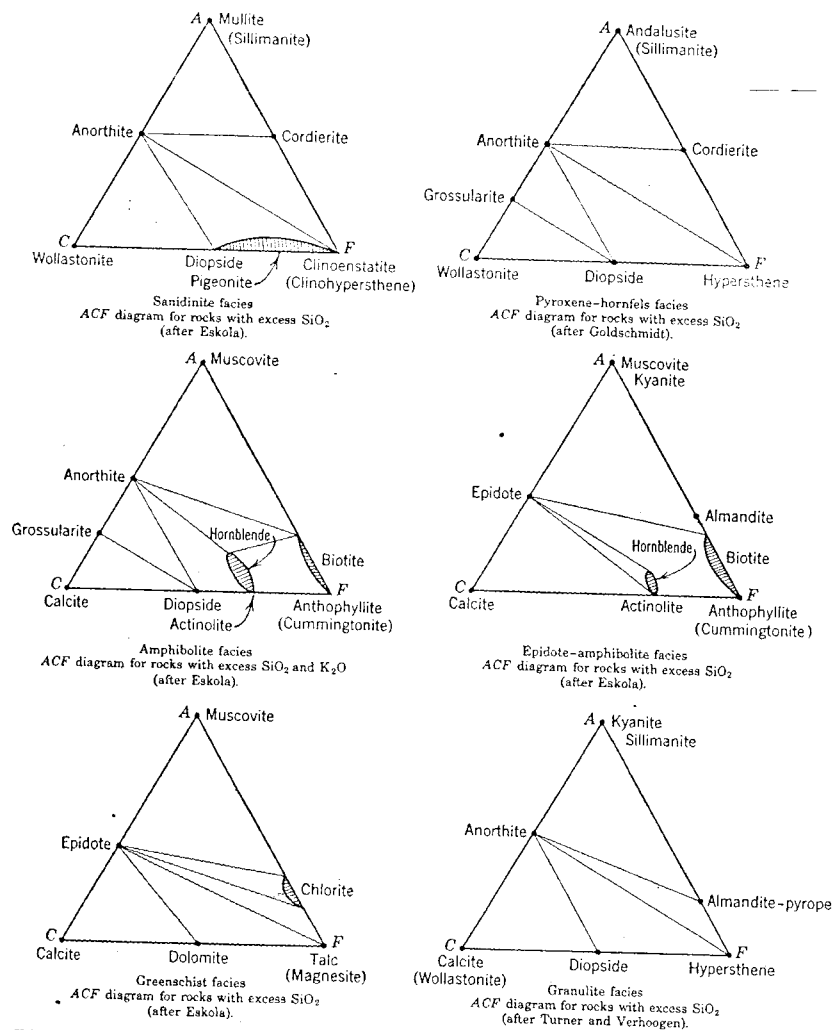
Figure 11.7   *ACF* diagrams for different facies.

can decompose silicates under these conditions, forming associations of quartz and carbonates.

The eclogite facies is characterized by pyroxenes of the omphacite type (intermediate between diopside and jadeite) and garnet with a high pyrope content. Other minerals that may be present are kyanite and hypersthene. It is difficult to draw a satisfactory *ACF* diagram for the eclogite facies, since known rocks belonging here have a comparatively small range in composition.

In the glaucophane-schist facies the critical minerals are glaucophane and lawsonite. Muscovite, almandine, epidote, and pumpellyite have also been recorded; these minerals are found in rocks of the greenschist and epidote-amphibolite facies, and glaucophane schists occur in close association with such rocks. This suggests that the

glaucophane-schist facies is related to these facies, but probably represents higher pressures, since glaucophane plus lawsonite can be considered as the high-density equivalents of plagioclase plus chlorite.

MINERAL TRANSFORMATIONS AND THE FACIES PRINCIPLE

Much remains to be done to make the facies classification the precise tool that it potentially is. This demands careful study of mineral assemblages and especially the determination of the variability in chemical composition of many of the characteristic minerals. The variability is linked not only with the bulk composition of the rock, but also with the probable temperature and pressure conditions under which it was formed. A mineral facies represents a definite (P, T) interval within which the rock attained its present mineralogical composition. The different facies are characterized by the stability fields of their critical minerals or mineral associations. The boundaries of a facies are delimited by the equilibrium curves for the reactions which produce the critical minerals and mineral associations.

In 1940 Bowen showed that the progressive metamorphism of limestones and dolomites could be considered in terms of a series of reactions of increasing decarbonation taking place at successively higher temperatures at any given pressure. He pointed out that the equilibrium curves for these reactions, when plotted on a pressure-temperature diagram, delimited fields of stability of specific minerals and mineral associations, thus forming what he termed a *petrogenetic grid*. At that time the equilibrium curves for the reactions he considered had not been determined, but he could arrive at approximations for them by evaluating the geological information provided by metamorphosed limestones and dolomites.

The thermodynamic basis for this petrogenetic grid and its relation to the facies principle have been thoroughly discussed by Thompson (1955), who points out that the boundaries between facies are curves of univariant equilibrium. If the equilibria involve only crystalline solids, then these curves have slopes determined by the Clapeyron equation and are virtually straight lines. Many significant equilibria in metamorphism, however, involve gain or loss of a volatile constituent such as carbon dioxide or water, acted on by a pressure equal to or less than that acting on the crystalline solids. If only one volatile constituent is involved in the equilibrium and the pressure on it is equal to the total pressure, the Clapeyron equation still applies, but the curves are concave toward the pressure axis and have a positive slope (since increasing pressure raises the decomposition temperature of a substance containing a volatile component).

These principles are illustrated in Figure 11.8, which gives equilibrium curves for a number of reactions pertinent to conditions of metamorphism. These curves have been determined either by experimental investigation of the equilibrium or by calculation of the equilibrium curve from thermochemical data. It can be seen that for reactions involving volatiles the curvature of the lines is greatest at low pressures, and they become virtually straight at comparatively moderate pressures.

As mentioned above, these curves represent the condition obtained when the pressure of the volatile component is equal to the total pressure. This condition may not
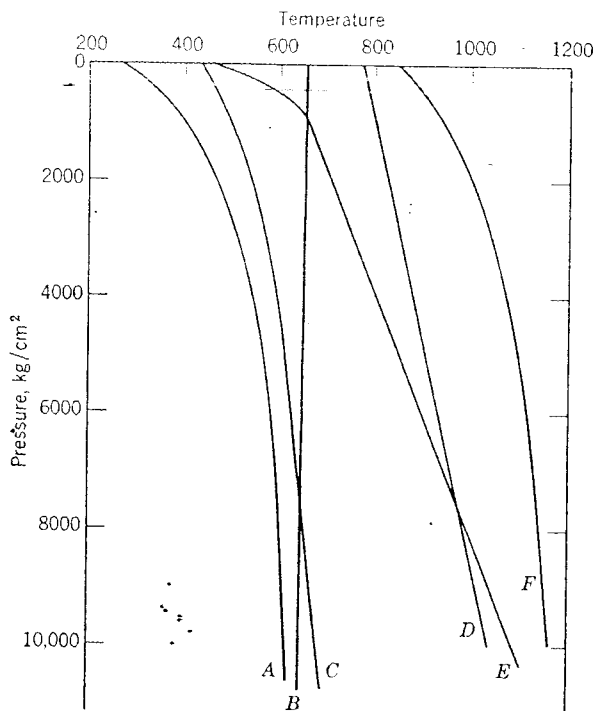
Figure 11.8 Equilibrium curves for some reactions of significance in metamorphism.

	Low Temperature	High Temperature
A.	Dolomite + quartz	= Diopside + carbon dioxide
B.	Orthoclase + albite	= Sanidine
C.	Pyrophyllite + corundum	= Sillimanite + quartz + water
D.	Almandine	= Hercynite + iron cordierite + fayalite
E.	Muscovite	= Sanidine + corundum + water
F.	Phlogopite	= Forsterite + leucite + $KAlSiO_4$ + water

apply in a specific case of metamorphism. For example, if the rock undergoing metamorphism is connected by pore spaces to the surface, the volatile constituent can leak away; its partial pressure may be very low and in the extreme case may approach that at the surface. Under these circumstances the equilibrium curve has a negative slope. This situation is illustrated by the calcite-wollastonite reaction (Figure 11.2). In effect, we have two limiting conditions for any reaction involving volatiles, and the temperature at which the reaction takes place depends on the extent to which the volatile component is able to escape.

The commonest volatile component in metamorphic reactions is water, and its significance in this connection has been emphasized by Yoder (1955). He points out that many discussions of metamorphism have tacitly assumed that water as a component is always present in sufficient amount to give rise to the most highly hydrated phases stable under the specific temperature and pressure conditions. Anhydrous

minerals having hydrous equivalents are then considered to indicate temperatures of formation above the equilibrium curve for reaction to give the hydrous equivalent. Forsterite, for example, reacts with water to give serpentine at about 500°. However, if water is not available, forsterite is stable below 500°, and in the absence of water it can form below 400° from magnesite and quartz. The logical consequence is that the mineralogy of a metamorphic rock depends not only on the composition of the original material and the temperature and pressure of metamorphism, but also on the amount and partial pressure of water and other volatiles. In terms of the facies principle we should be able to distinguish water-deficient and water-sufficient variants of the same facies. For example, in rocks of the greenschist facies the usual calcium aluminosilicate is zoisite or epidote; however, in some tuffaceous graywackes that have suffered low-grade metamorphism prehnite and calcium zeolites, minerals of similar composition except for a higher water content, are found instead. Similarly, the amount and partial pressure of carbon dioxide have a marked effect on the mineralogy of rocks containing appreciable amounts of calcium and magnesium. In the amphibolite and granulite facies both wollastonite and calcite plus quartz are probably stable, their occurrence being conditioned by the partial pressure of carbon dioxide.

A noteworthy feature of Figure 11.8 is that the reaction curves are all steeply inclined to the temperature axis; in other words, they are much more strongly temperature-dependent than pressure-dependent. Such reactions are suitable for assessing temperatures of metamorphism but not for assessing pressure. There is a notable lack of reactions that are strongly pressure-dependent and can thus serve as *geological barometers* and indicators of depth of crystallization. It is known the magnesium-rich garnets characteristically occur in rocks that are believed to have crystallized at great depth, such as eclogites, but the equilibrium relations of such garnets are little known. Diamond is probably a good geological barometer, but its restricted occurrence limits its usefulness. Aluminum substitution in orthopyroxenes may serve as an indication of pressure conditions, but the situation is complicated if other aluminous minerals are present in the assemblage. One relationship that is strongly pressure-dependent is that between kyanite and the other polymorphs of Al_2SiO_5, sillimanite and andalusite, kyanite having a density about 3.6 and the other polymorphs about 3.2. Laboratory measurements confirm that the equilibrium is strongly pressure-dependent and that kyanite is the high-pressure phase at low to medium temperatures (Figure 11.1). Aragonite, the high-pressure phase of $CaCO_3$, has been found in jadeite-bearing metagraywackes of the Franciscan formation in California. Assuming a temperature of metamorphism of 200°, the minimum pressure of formation indicated by the presence of jadeite is 5 kb; aragonite would require a pressure of 7 kb. Under hydrostatic conditions these pressures imply a minimum depth of metamorphism for these rocks of 20 to 25 km.

We are handicapped in our elucidation of the geochemistry of metamorphism by our lack of knowledge of the equilibrium conditions of many important metamorphic reactions, especially those taking place at lower grades. The significant reactions are often very complex. For example, one of the commonest reactions of low-grade metamorphism is muscovite + chlorite = biotite, and the incoming of biotite is an impor-

tant marker in many regions of progressive metamorphism. However, rather than being a direct combination of muscovite and chlorite it may be a reaction between muscovite of one composition and chlorite to give muscovite of a different composition and biotite, or a reaction involving potash feldspar. The sequence of reactions is shown in Figure 11.9. At the lower temperature biotite is formed by the reaction microcline + chlorite → biotite + white mica + quartz + H_2O. At higher temperatures the change in muscovite composition outlined above takes place. The decreasing field of muscovite-phengite stability may be seen in the diagram. Phengite is the mineral name for a muscovite-type mineral with a substitution of (Mg, Fe^{2+}) Si for the AlAl in ideal muscovite.
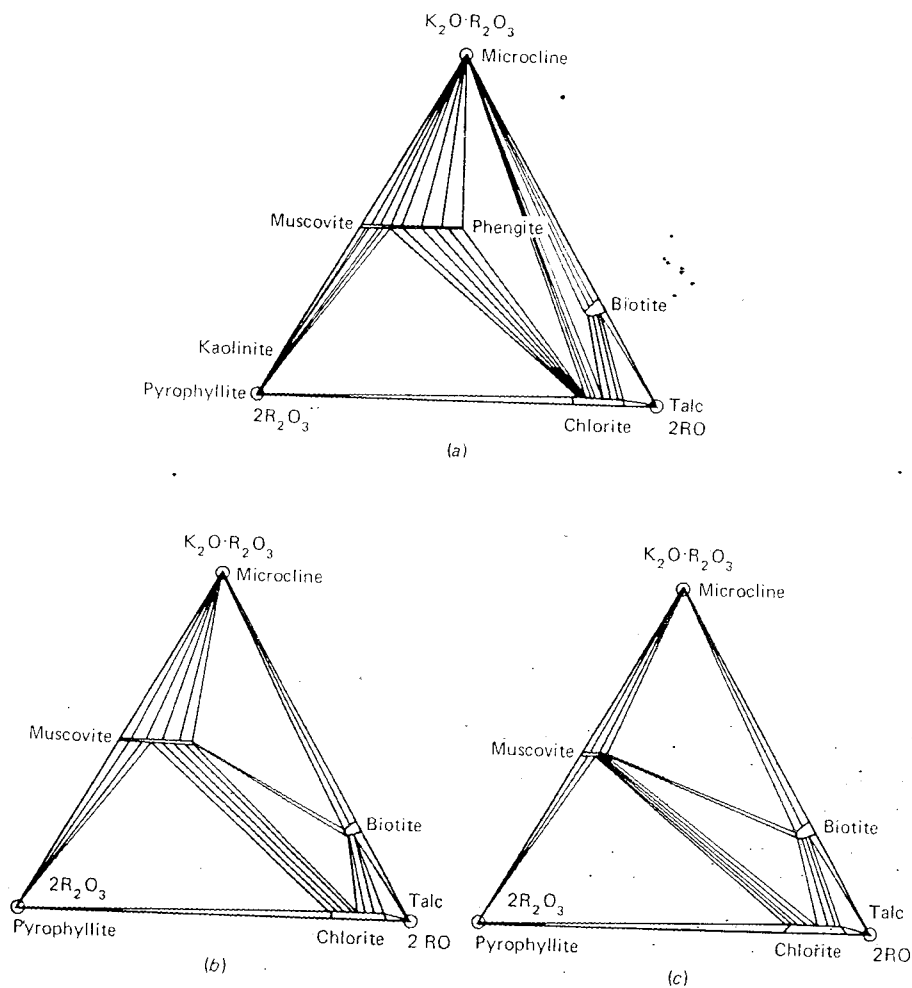
Figure 11.9 Schematic *AFM* diagrams representing variations of paragenetic relations of garnet and cordierite in metapelites with temperature and pressure. $A = Al_2O_3 - (K_2O + Na_2O)$, $F = FeO$, $M = MgO$. (Modified from Chinner, 1962.)

Most minerals of metamorphic rocks are capable of considerable variation in composition through atomic substitution, and the stability of a mineral and its reactions with other minerals will be affected by the nature and extent of these substitutions. For example, garnet in metamorphic rocks varies from iron-rich almandine in low-grade schists to magnesium-rich pyrope in high-grade eclogites and granulites. Similarly, the anorthite content of plagioclase increases with increasing metamorphic grade, as does the aluminum content of hornblende.

It has long been recognized that the anorthite content of plagioclase in metamorphic rocks often increases with increasing degree of metamorphism. This is a reflection of the independent role of sodium and calcium in these rocks. Sodium is generally present almost entirely in the form of the albite component of plagioclase, whereas calcium may be present in a number of calcium aluminum silicates. In most dynamically metamorphosed rocks of low-grade, albite is associated with zoisite or a mineral of the clinozoisite-epidote series. As the degree of metamorphism increases the plagioclase becomes more calcic and the amount of zoisite decreases. The zoisite is converted to anorthite, which enters the plagioclase. The reaction can be represented by the equation

$$Ca_2Al_3Si_2O_{12}(OH) = 3CaAl_2Si_2O_8 + CaO + H_2O$$
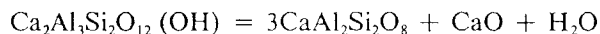
This is certainly an oversimplification, but it shows that the reaction involves water, and the equilibrium will be affected by the partial water-vapor pressure. If we consider the reaction to take place at some fixed water-vapor pressure, the relationship between degree of metamorphism and plagioclase composition can be represented diagrammatically (Figure 11.10). The relationship is additionally complicated by the peristerite unmixing in sodium-rich plagioclase (Chapter 5). Electron microprobe analyses have shown that plagioclase in rocks of the greenschist and epidote-amphibolite facies is almost pure albite ($Ab_{>99}$), whereas in rocks of the amphibolite facies it is Ab_{80} or more calcic. This is illustrated in Figure 11.10, which shows that albite and zoisite (epidote) are in stable equilibrium up to temperature t_2. At this temperature the albite is entirely converted to plagioclase of composition Ab_{80}, provided
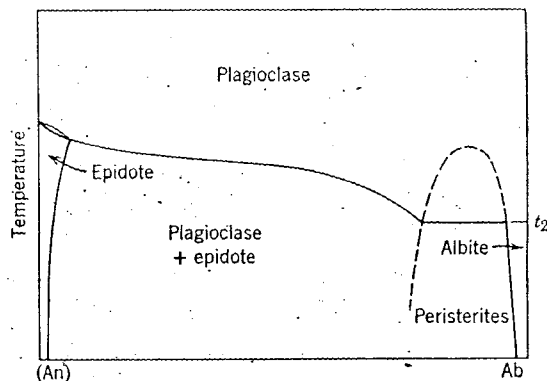
Figure 11.10   Stability relations between zoisite (epidote) and plagioclase. (After Christie, Norsk Geol. Tidssk. 39, 270, 1959)

sufficient calcium is available; insufficient calcium results in a mixture of two plagioclases, about Ab_{99} and Ab_{80} respectively. As the temperature increases beyond t_2, if calcium is still available, the plagioclase becomes progressively richer in the anorthite component.

Transformations of the type discussed above can thus be used as recorders of the conditions under which metamorphism took place. Careful study of such transformtions, both in laboratory preparations and as revealed in metamorphic rocks, can contribute greatly to a more precise knowledge of the different facies and their mutual relationship.

METASOMATISM IN METAMORPHISM

In discussing the facies principle we have considered metamorphic rocks as the end-products of chemical reaction; the equilibrium assemblage of minerals has been determined by the conditions of metamorphism and the ultimate composition attained by the rock in question. We have not been concerned with the possible changes in composition that the rock may have undergone during metamorphism, that is, whether metasomatism has played a part. It remains to consider some of the problems posed by metasomatism, apart from the mode of transportation of the metasomatizing substances, which was discussed in a previous section.

The reactions accompanying metasomatism are governed, of course, by physico-chemical principles, of which the phase rule and the law of mass action are the most significant in this connection. However, an additional rule has been found to apply to metasomatism: experience has shown that metasomatism usually takes place without change of total volume; that is, the metasomatized rock generally occupies the same volume as it did before metasomatism. This is highly significant in that metasomatic reactions are not quantitatively represented by the conventional simple equation, which merely balances equal weight of material on the right and left sides, respectively, without taking volume changes into account. Thus the serpentinization of olivine can be represented by the following equations:

$$3Mg_2SiO_4 + SiO_2 + 4H_2O = 2H_4Mg_3Si_2O_9$$

or

$$5Mg_2SiO_4 + 4H_2O = 2H_4Mg_3Si_2O_9 + 4MgO + SiO_2$$
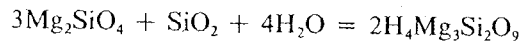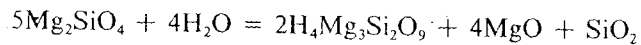
Of these two equations the second better approximates the conversion of olivine into an equal volume of serpentine (assuming excess MgO and SiO_2 are removed in solution) and may be expected to represent more closely the actual process. It is thus unsafe to interpret metasomatism quantitatively in terms of simple equations that actually correspond to reactions involving marked charges of volume, although such equations are useful for indicating qualitatively the general direction in which chemical reaction may have proceeded.

Some instances of metasomatism have already been mentioned, of which the formation of dolomite by the action of magnesium-bearing solutions on calcite, discussed in Chapter 6, is one. Many ore deposits seem to represent a metasomatism of pre-existing rocks, the ore material having replaced them volume for volume. Meta-

somatism very often accompanies the late stages of magmatism. Many china clay deposits, such as those of Cornwall, are evidently the result of the metasomatic alteration of granite, whereby the feldspar is altered to kaolin. Schematically

$$2KAlSi_3O_8 + 2H_2O = Al_2Si_2O_5(OH)_4 + K_2O + 4SiO_2$$

That the metasomatism takes place volume for volume is shown by the way in which the kaolin occurs as pseudomorphs after the feldspar. Ultrabasic rocks are particularly subject to metasomatism. Conversion to serpentine is the commonest form of alteration, but if calcium is present or introduced, assemblages such as actinolite-chlorite may form. At temperatures corresponding to the greenschist facies CO_2 may replace SiO_2, and talc-dolomite, dolomite-quartz, or magnesite-quartz rock may be the ultimate products. Some of the best-described examples of metasomatism are those accompanying ore deposits, since they are often well exposed by mining operations and their interpretation is important for an understanding of the economic geology. They were first described in connection with ore deposits in the strongly metamorphosed Precambrian of Sweden, and the Swedish term *skarn* has become a general term for such material. Extensive skarn masses generally surround iron ores in limestones; the major skarn minerals are usually garnet (andradite), pyroxene (hedenbergite), and amphibole (actinolite and hornblende). Often the source of the metasomatizing solutions can be traced to neighboring intrusions of granite; sometimes the skarns are attributed to interchange of material between the ore body and the country rock during regional metamorphism.

An approach to the quantitative study of metasomatic change has been developed by Helgeson (1967) using equilibrium and nonequilibrium thermodynamics. The approach used may be illustrated using an activity diagram as shown in Figure 11.11. Activity diagrams of this type were devised by the geochemist Korzhinski and are often called Korzhinski diagrams. Figure 11.11 illustrates the stable mineral phases present in the system MgO—K_2O—Al_2O_3—SiO_2—H_2O at 300° C. Other diagrams may be constructed to extend the system to include CaO and Na_2O, and to approximate sea water. Point A in the diagram represents the composition of sea water and may be considered for this argument as typical of connate water in pelitic sediments. The dashed lines illustrate the path of chemical change as sea water reacts with K-feldspar in a metamorphic system. The initial reaction will be according to the equation

$$2KAlSi_3O_8 + 8H_2O + 5Mg^{2+} \rightarrow Mg_5Al_2Si_3O_{10}(OH)_8 + 8H^+ + 2K^+ + 3SiO_2$$
  K-feldspar   chlorite   quartz

As the reaction proceeds the pH of the solution decreases until point B is reached where kaolinite begins to form along with chlorite and quartz. With the simultaneous formation of chlorite and kaolinite the activity ratio of Mg^{2+}/H^+ remains constant but the pH increases because H^+ is utilized in the formation of kaolinite. At C kaolinite reacts with K-feldspar to form muscovite according to the reaction

$$KAlSi_3O_8 + Al_2Si_2O_5(OH)_4 \rightarrow KAl_3Si_3O_{10}(OH)_2 + 2SiO_2 + H_2O$$
    K-feldspar   kaolinite   muscovite   quartz

(a) Plus quartz, calcite, and solution

(b) Plus quartz, calcite, chlorite, and solution

Figure 11.11   Activity diagrams depicting phase relations in the Salton Sea geothermal system at 300°C. The solid round symbol represents the composition of the geothermal formation fluid and the solid and open square symbols at *A* and *J* respectively represent sea water (Garrels and Thompson, 1962) and lower Colorado River water (Livingstone, 1963). The dashed arrows and letter annotations are explained in the text along with the basis for placing the phase boundaries in the positions shown. The activity of water was taken as unity in constructing these diagrams. (Helgeson, 1967)

When all the kaolinite is consumed the reaction moves along path *CD* to *D* where equilibrium has been established between solution muscovite, quartz, and K-feldspar. A calculation of the net mass transfer of material indicates the solution of 3.996 g $K^+/10^3$ g H_2O and the loss of 1.244 g $Mg^{2+}/10^3$ g H_2O. Thus simple equilibration reactions of mineral with sea water may make K^+ available for subsequent metasomatic reactions as H_2O solutions are driven from the pelites by metamorphism. In predicting metamorphic reactions many alternative systems and paths exist and require some estimation of temperature of metamorphism and the ease with which water is removed.

The nature of metasomatic reactions on a regional scale has been debated for a long time. It does appear that major elemental changes take place by the action of

water solutions and hence will be most pronounced at lower temperatures of metamorphism when water is a significant component. Element migration at high temperatures becomes significant when the partial melting temperatures of metamorphic rocks are reached.

ACCESSORY ELEMENTS IN METAMORPHIC ROCKS

Whereas considerable data on the abundance and distribution of minor and trace elements are available for igneous rocks and, to a lesser extent, for sedimentary rocks, corresponding information for metamorphic rocks is conspicuously scanty. Variety of metamorphic rock types and difficulties in making generalizations contribute to this lack of meaningful data. Likewise metamorphic minerals are most often not homogeneous and sampling therefore difficult. For example, many garnets show significant zoning due to the low rate of diffusion in them; only the outer rims are in equilibrium with the surrounding rocks. This has led to much speculation on patterns of trace element behavior and their possible geological significance in metamorphic terranes. For example, the occurrence of tourmaline in schists derived from sedimentary rocks has been cited as evidence for boron metasomatism from supposedly deeper-lying granite intrusions; however, since many marine clays show a notable enrichment in boron, this may be sufficient to account for the tourmaline in many schists. Theoretical considerations enable some predictions as to possible patterns of behavior for certain elements. Strongly lithophile trace elements may be expected to follow Goldschidt's rules for incorporation in the crystal structures of the common minerals of metamorphic rocks. The greater variety of minerals and of conditions of crystallization in metamorphic rocks, compared to igneous rocks, will result in more complex relations. The greater possibilities of disorder in crystal lattices at higher temperatures suggest greater tolerance for foreign elements in minerals formed at higher grades of metamorphism. For trace elements of chalcophile or weakly lithophile character electronegativity and the concentration of sulfide ions are probably more significant than ionic size in determining their ultimate deposition.

Under ideal conditions the distribution of accessory elements between metamorphic minerals is governed by thermodynamic parameters. The regularities of thermodynamic equilibrium allow the potential use of trace elements as indicators of the temperatures and pressures of metamorphic reactions. The distribution of a trace element between two phases is related by the Nernst partition law. In the case of trace elements, which behave as dilute solutions, the relationship

$$K = \frac{X_\alpha^A}{X_\alpha^B}$$

holds. K, the distribution coefficient, is a function of the concentration X of trace element α in mineral phases A and B.

The distribution coefficient varies with temperature according to the relations

$$\left(\frac{\delta \ln K}{\delta T}\right)_P = \frac{\Delta H}{RT^2} \quad \text{and} \quad \left(\frac{\delta \ln K}{\delta P}\right)_T = \frac{\Delta V}{RT}$$

Studies by Kretz (1959) have shown some metamorphic systems in which these relationships fit very nicely and others in which the number of variables is so great that significant conclusions concerning T and P are masked by major composition variations. The most useful distribution studies in metamorphic systems appear to be those of oxygen isotopes, and temperature estimates for metamorphic facies are generally based on such studies.

One of the few investigations of minor and trace elements in metamorphic rocks for which both geological and geochemical data are closely controlled is that of Shaw (1954). He selected the Littleton formation of New Hampshire for examination, because it is a series of pelitic rocks that has undergone progressive regional metamorphism, and its present condition shows a transition from shales through increasingly metamorphosed rocks to sillimanite schists and gneisses. Analyses were made of 63 samples of the formation, representing all grades of metamorphism. The results of this extensive study showed that the concentration of most elements remained constant during the metamorphism. The only detectable changes in trace elements were a slight decrease in nickel and copper and a well-defined increase in lithium and lead, the latter being correlated with some degree of potassium metasomatism.

This work of Shaw has general significance because it shows that a formation of fairly uniform composition retains the primary pattern of minor and trace elements even after extensive and intensive metamorphism. Thus the pattern of minor and trace elements may be a useful guide to the character of the original rock. This principle has been applied to the problem of the original nature of certain amphibolites. Nearly identical amphibolites, consisting essentially of hornblende and plagioclase, may evolve from skarnlike metasomatism of carbonate sediments or from essentially isochemical metamorphism of gabbroic and dioritic rocks. In such amphibolites the most satisfactory clue to their origin may lie in their distinctive inheritance of accessory elements. Dynamothermal metamorphism of gabbroic rocks does not generally remove the initial concentrations of such elements as chromium, cobalt, nickel, and copper; conversely, amphibolites derived by the replacement of marbles tend to be deficient in these elements and show higher concentrations of barium and lead.

Under some circumstances, however, it appears that appreciable changes in pattern of accessory elements are possible. DeVore (1955) has made a large number of analyses of the individual minerals extracted from rocks of different metamorphic facies, and claims that the replacement of one mineral assemblage by another will be accompanied by a redistribution and fractionation of both major and minor constituents. For example, he suggests that the transformation of an epidote-amphibolite facies hornblendite to a granulite facies hornblendite could release large amounts of Cr, Ni, Cu, and Mg, and the reverse transformation could release Pb, Zn, Ti, Mn, and Fe. His data indicate that a cubic kilometer of epidote-amphibolite hornblendite changing to granulite facies hornblendite could release eight million tons Cr_2O_3, four million tons of NiO, and 800,000 tons of CuO, and the reverse transformation could release 11 million tons MnO, 94 million tons TiO_2, 800,000 tons ZnO, and 27,000 tons PbO. This would provide an excellent source of ore-forming elements, but no convincing mechanism for the extraction of this material and its concentration into

ore bodies has been proposed, nor have instances of ore deposits formed in this way been described. Under these circumstances this concept must be considered as an interesting speculation rather than an established fact.

ULTRAMETAMORPHISM

If the temperature continues to rise during metamorphism, any rock must eventually melt. In this way a magma is generated, and the further geochemical evolution is no longer part of metamorphism. The regeneration of magma will not, however, take place at a definite temperature and pressure but over a range of temperatures and pressures; the process may not proceed to completion and may be halted at any stage. Mixed rocks are thereby formed with characters partaking of those of both igneous and metamorphic rocks.

The process can be looked upon as the reverse of magmatic crystallization. In the melting of rocks we may expect the first liquid formed to resemble the last liquid fraction of a magma and be rich in silica, soda, potash, alumina, and water. The process of differential fusion or *anatexis* can be conceived as beginning with the "sweating" of the low-melting fraction from the main mass of the rock and its segregation into lenses. In the light of laboratory work on the crystallization of hydrous feldspar melts differential fusion of rocks may be expected to begin at temperatures of 600–700°. If the process ceases at this point, the fused material will crystallize as an aggregate of quartz and feldspar as lenses within the more refractory material, giving rise to the rock type that has been called venite, a particular variety of migmatite. A similar product could be formed by the injection of the last fraction from magmatic crystallization between the layers of a solid rock (lit-par-lit injection); the resulting rock has been called arterite to distinguish it from one formed by differential fusion. In the field it is often difficult to decide whether a migmatite is an arterite or a venite, and indeed both processes outlined above may have been active at the time of formation.

Under conditions of ultrametamorphism the regenerated magma will most likely be granitic, because of the abundance of quartz-feldspathic sediments and the lower melting temperatures of hydrous rocks with granitic composition. Some granites may well have been formed by the actual remelting of material of suitable composition. However, granites may also originate without remelting. The mineral association of granite—quartz, potash feldspar, biotite and/or hornblende—is typical of the amphibolite facies, and any body of rock with a bulk composition corresponding to that of granite (or which is coverted thereto by metasomatism) recrystallizes to give this typical mineral association under conditions of that facies. Generally, as a result of directed pressure during crystallization, the product is a gneiss, but, if directed pressure is weak or lacking, a normal granite may result. Thus granites may be produced in several ways: by the fractional crystallization of a magma; by the crystallization of a melt produced by the differential fusion of a preexisting rock; and by the recrystallization without fusion of a preexisting rock. Any body of granite may include within it representatives of all these types of origin. Deciding the mode of formation of a particular granite may require all the resources, field and laboratory, of a geologist, and even so, the answer may be equivocal. The volume of discussion

on the origin of granite is an eloquent expression of the difficulty in determining the boundary between magmatic and metamorphic processes in actual rocks.

SELECTED REFERENCES

Barth, T. F. W. (1962). *Theoretical petrology* (second ed.). 416 pp. John Wiley & Sons, New York. A comprehensive and logical presentation of principles and underlying theory of the petrology of metamorphic rocks.

Bowen, N. L. (1940). Progressive metamorphism of siliceous limestone and dolomite. *J. Geol.* **48**, 225–274. A classic paper in which the concept of a petrogenetic grid was first formulated.

Burnham, C. W. (1959). Contact metamorphism of magnesian limestones at Crestmore, California. *Geol. Soc. Amer. Bull.* **70**, 879–920.

DeVore, G. W. (1955). The role of adsorption in the fractionation and distribution of elements. *J. Geol.* **63**, 159–190. Discusses the changes in trace-element content between metamorphic facies and the possibility that metamorphic transformations liberate ore-forming materials.

Ernst, W. G. (ed.) (1975). *Metamorphism and plate tectonics regimes.* 445 pp. Dowden, Hutchison & Ross Inc. Stroudsburg, PA. Twenty-two papers on metamorphism and metamorphic belts.

Ernst, W. G. (1976). *Petrologic phase equilibria,* 333 pp. W. H. Freeman and Co. San Francisco. Includes diagrams related to metamorphic rocks.

Garlick, G. D. and S. Epstein (1967). Oxygen isotope ratios in coexisting minerals of regionally metamorphosed rocks. *Geochim. Cosmochim. Acta* **31**, 181–214. The estimation of temperatures of regional metamorphism by oxygen isotope studies.

Helgeson, H. C. (1967). Solution chemistry and metamorphism. In P. H. Abelson (ed.) *Researches in geochemistry,* Vol. 2. John Wiley & Sons, New York.

Hofman, A. W., B. Giletti, H. S. Yoder, Jr., and R. A. Yund (eds.) (1974). *Geochemical transport and kinetics.* 352 pp. Carnegie Inst. of Washington, Washington, D.C. A collection of papers on diffusion and reaction kinetics in rocks.

Kretz, R. (1959). Chemical study of garnet, biotite and hornblende from gneisses of southwestern Quebec, with emphasis on distribution of elements in coexisting minerals. *J. Geol.* **67**, 371. An example of the distribution of trace elements in metamorphic rocks and their interpretation.

Miyashiro, A. (1972) Metamorphism and related magmatism in plate tectonics. *Amer. J. Sci.* **272**, 629–656. Relates plate tectonics to igneous and metamorphic rocks.

——— (1973). *Metamorphism and metamorphic belts.* 492 pp. Halsted Press, John Wiley, New York. A classic reference source on metamorphism.

Packham, G. H. and K. A. W. Crook (1960). The principle of diagenetic facies and some of its implications. *J. Geol.* **68**, 392–407. An illuminating discussion of the relationship of diagenesis to metamorphism.

Ramberg, H. (1952). *The origin of metamorphic and metasomatic rocks.* 317 pp. University of Chicago Press, Chicago. A comprehensive account of metamorphism and metamorphic rocks, with emphasis on their interpretation in terms of thermodynamics and crystal chemistry.

Shaw, D. M. (1954). Trace elements in pelitic rocks. *Bull. Geol. Soc. Amer.* **65**, 1151–1152. Gives a statistical analysis of the content of minor and trace elements in 63 rocks from the same formation, representing all grades of metamorphism from shales to sillimanite gneisses.

Thompson, J. B. (1955). The thermodynamic basis for the mineral facies concept. *Amer. J. Sci.* **253**, 65–103. A thorough treatment of the thermodynamics of metamorphism and the geological applications.

Turner, F. J. (1981). *Metamorphic Petrology* 2nd ed. 524 pp. McGraw-Hill, New York. Mineralogical, field and tectonic aspects of metamorphic rocks are discussed.

Vernon, R. H. (1975). *Metamorphic Processes.* 247 pp. John Wiley & Sons, New York. A useful review of metamorphic processes.

Winkler, H. G. F. (1979). *Petrogenesis of metamorphic rocks* (fifth ed.) 348 pp. Springer-Verlag, New York. A thorough discussion of the chemical and mineralogical transformations accompanying rock metamorphism.

Yoder, H. S. (1955). Role of water in metamorphism. *Geol. Soc. Amer. Spec. Paper* **62**, 505–524. Criticizes the concept of the ubiquitous occurrence and adequate supply of water in metamorphism and discusses the consequences of a water-deficient environment.

Yoder, H. S. and H. P. Eugster (1955). Synthetic and natural muscovites. *Geochim. Cosmochim. Acta* **8**, 225–280. An exhaustive study of one of the most abundant and widespread metamorphic minerals.

Zen, E-an and A. B. Thompson (1974). Low-grade regional metamorphism. Mineral equilibrium relations. *Ann. Rev. Earth Planet. Sci.* **2**, 179–212. A review of low-grade metamorphism especially of the zeolite facies. Extensive bibliography.

12 THE GEOCHEMICAL CYCLE

THE EARTH AS A PHYSICOCHEMICAL SYSTEM

Geochemically, the earth may be considered a closed system, as this term is used in physical chemistry. This concept can be criticized as an oversimplification; we have seen that some material—meteorites and meteoritic dust—is continually being received from outer space, and that some hydrogen and helium is being lost by escape from the upper atmosphere. Nevertheless, these gains and losses are insignificant as compared with the system as a whole. If our interpretation of the geological record is correct, the chemical processes taking place on the surface and within the earth have probably operated with a remarkable degree of uniformity for the last 4000 million years.

By the nature of things the geochemist is mainly concerned with the surface of the earth, since it is the only part accessible to direct examination. In the discussion of the origin and structure of the earth an attempt was made to present a logical account of the probable nature of the interior. The data on which this account was based are indirect, their interpretation speculative. However, the general picture of a nickel-iron core, a mantle largely of magnesium-iron silicate, and a crust in which oxygen, silicon, aluminum, iron, calcium, magnesium, sodium, and potassium are the major constituents gives a consistent interpretation of information gleaned from many independent sources—the study of meteorites, the physics and chemistry of the earth, seismological data, and so on. We may therefore accept such a picture as a working hypothesis, realizing always that it is a hypothesis, but a well-buttressed one.

One of the few subjects upon which universal agreement seems to prevail is that the earth was not created in its present state. The geochemical evidence supports the

idea that its internal structure is probably the result of forces originating within the earth itself. The earth is a system with considerable mass and thus exerts a gravitational force on its own components. The resulting gravitational field has affected the distribution of material by concentrating the heavier phases toward the center and the lighter toward the surface. The rate of such a gravitational differentiation clearly depends on the viscosity of the system; it is more rapid in a gas than in a liquid, more rapid in a liquid than in a solid. The idea of an earth at one time liquid is attractive, for it explains how such gravitational differentiation could take place within a comparatively short time. Geophysical data sustain the theory that the earth has a layered structure due to separation of its material into shells of different density. Goldschmidt termed this the primary geochemical differentiation. It was a differentiation due to gravity acting on a system in which iron, oxygen, magnesium, and silicon were the major components. Iron was the principal component, and the distribution of the elements between a metallic core and a silicate mantle was controlled by their oxidation potential with respect to that of iron. Elements more readily oxidized than iron concentrated in the mantle; the others alloyed with iron to form the core. Hence the fate of an element in this primary geochemical differentiation is in effect a reflection of the number and arrangement of its orbital electrons. Those elements forming ions with a noble gas structure went into the silicate phase; the transition elements, on the other hand, concentrated in the metallic core or in a sulfide phase.

THE CRUST AS A SEPARATE SYSTEM

The separation of crust, mantle, and core enables us to consider the outer part of the earth as a distinct physicochemical system. The concepts of global tectonics indicate that the crust interacts with a portion of the upper mantle. Seismic evidence shows that the asthenosphere in the depth range of 75–250 km is plastic or partially molten. This plastic region allows the movement of lithospheric plates. Since this region is partially molten it may chemically interact with layers above it, especially at divergent or convergent plate margins. At divergent margins new material is added to the earth's surface. At convergent margins subduction of surface materials into the mantle allows them to be melted or mobilized and recycled to the earth's surface (Figure 12.1). The migration of material within the crust and upper mantle can then be discussed as an independent phenomenon; it is partly mechanical, brought about by orogenic movements or gravitational forces, and partly chemical. Mechanical movements belong in the field of geology. Geochemistry is concerned with the migration of the elements under the influence of physicochemical forces. This migration has been discussed in terms of the processes of magmatism, sedimentation, and metamorphism. We have seen that the fate of an element during magmatic crystallization is primarily a function of its ionic size. A particular element appears in those minerals in the lattices of which it fits most readily and with the greatest decrease of free energy. The distribution of the elements by ionic size in this way was described by Goldschmidt as the secondary geochemical differentiation. Magmatic crystallization also adds important amounts of a few elements to the atmosphere and hydrosphere.
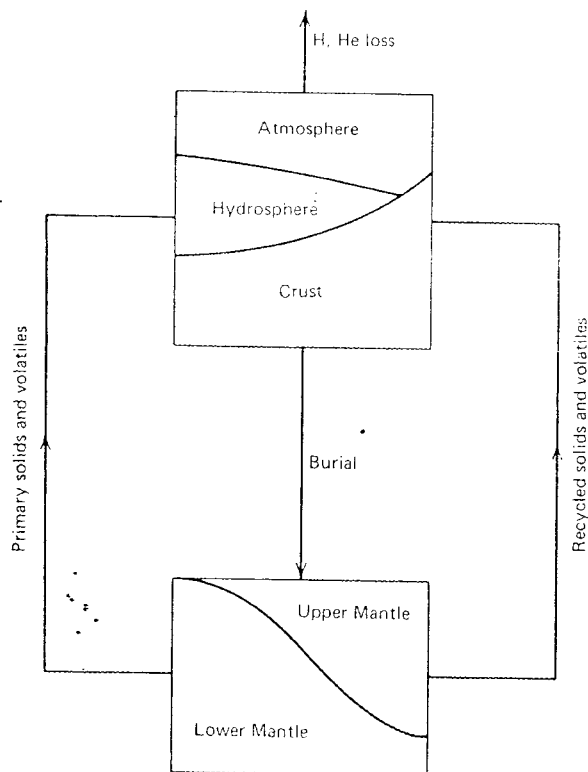
Figure 12.1 Schematic model of major long-term geo-
chemical processes at the earth's surface.

The processes of sedimentation can be looked upon as leading to a further degree
of geochemical differentiation. The temperature and pressure conditions at the
earth's surface that permit a hydrosphere to exist provide a chemical environment in
which extensive differentiation may take place. Sedimentary and hydrospheric pro-
cesses provide the mechanism for major geochemical differentiation processes. At the
comparatively low temperature of sedimentary processes ionic substitution in min-
erals is much less prominent, although still significant. Coprecipitation under a par-
ticular set of physicochemical conditions is, however, an important way in which
certain elements become associated in specific types of sediment. The controlling
factors are those pertaining to the properties of ions in aqueous solutions, and the
ionic potential is of primary importance. Geochemical differentiation during sedi-
mentation is therefore governed in large part by interrelationship between ionic
radius and ionic charge.

Sedimentation also involves an interaction of the hydrosphere and the atmosphere
with the lithosphere. Water and carbon dioxide influence the weathering of rocks
and minerals and provide the buffer system for pH in the hydrosphere. A schematic
diagram illustrating the interaction of hydrosphere, atmosphere, and sediments—the
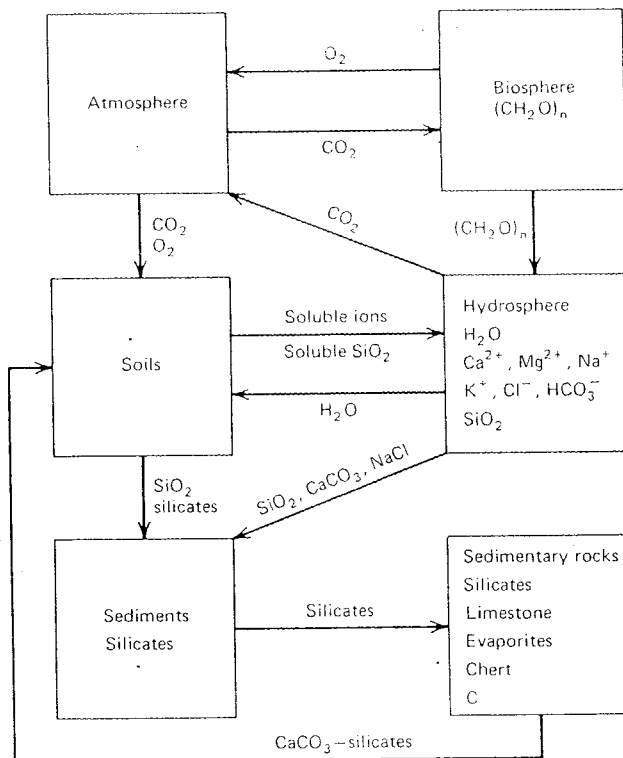
Figure 12.2   Exogenic or surface geochemical cycle.

exogenic cycle—is shown in Figure 12.2. Water and carbon dioxide are incorporated into sedimentary minerals, soluble ions (especially Na^+) are contributed to the hydrosphere, and SiO_2 is liberated.

· Processes involving living organisms are intimately associated with sedimentation but can be considered separately, and these processes are even more closely linked with the hydrosphere and the atmosphere. We have seen how photosynthesis has probably been largely responsible for the present composition of the atmosphere and how the balance of dissolved material in the ocean is largely a function of the organic life therein. Thus in the biosphere a further geochemical differentiation takes place through the metabolic action of organisms.

The series of changes so far discussed has led on the whole to an increasing degree of geochemical differentiation. This tendency is reversed by metamorphism. In general, metamorphism tends toward uniformity of distribution of the elements. One can visualize unlimited metamorphism as resulting in an ideal condition in which the whole lithosphere reaches a uniform composition. This may seem an exaggerated view, but that such a tendency exists is evidenced by the comparatively monotonous chemical and mineralogical composition of ancient geological formations. Compared with the chemically diverse rock types of younger ages, the Archaean is dominantly

made up of gneisses of relative uniformity, probably due in considerable degree to long-continued metamorphic and metasomatic reactions.

Thus in the beginning the relative abundances of the elements were determined by the stability of their isotopes, that is, by the number of neutrons and protons in the nucleus and the binding energy. The processes by which the earth was formed led to a first separation of elements according to their volatility or ability to form volatile compounds: the earth is evidently impoverished in the more volatile elements and compounds in comparison with the universe as a whole. The separation of the earth into an iron core and a silicate mantle and crust resulted in a strong fractionation of the elements according to their affinity for metallic iron or for silicate; this fractionation was controlled by the number and arrangement of the outer electrons. The next step in the evolution of the earth was the solidification of mantle and crust, which led to a further fractionation, this time determined largely by the role of the different elements in liquid $\rightleftharpoons$ crystal equilibria. The major controlling factor was ionic size. During geological time considerable fractionation of the elements has taken place at the earth's surface as a result of sedimentary processes; the fate of an element under these conditions is largely a matter of its ionic potential, the ratio of ionic size to ionic charge. The absolute abundance of an element is conditioned by its nuclear structure; its abundance in a particular part of the universe or of the earth is conditioned by more superficial atomic characters, such as the number and arrangement of the orbital electrons, and the size of the atom or ion. The geochemical behavior of each element depends on its individual properties under the physicochemical conditions at each stage in the geochemical cycle. Figure 12.3 summarizes the varying physicochemical conditions under which principal geochemical reactions now take place.
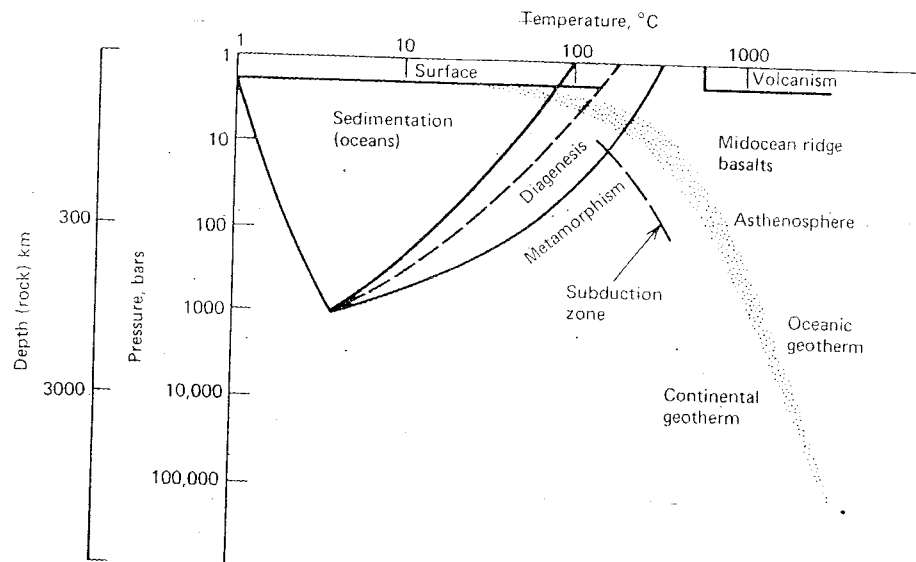
Figure 12.3 Zones of important geochemical reactions.

THE GEOCHEMICAL CYCLE

This overall picture of the migration of the elements in the outer part of the earth provides us with the concept of the geochemical cycle (Figure 12.4). In the lithosphere the geochemical cycle begins with the initial crystallization of a magma, proceeds through the alteration and weathering of the igneous rock and the transportation and deposition of the material thus produced, and continues through diagenesis and lithification to metamorphism of successively higher grade until eventually, by anatexis and palingenesis, magma is regenerated. Like any ideal cycle, the geochemical cycle may not be realized in practice; at some stage it may be indefinitely halted or short-circuited or its direction reversed. The geochemical cycle is not closed, either materially or energetically. It receives "primary" magma from below bringing energy with it in the form of heat. The surface receives an insignificant contribution of meteoritic matter, which nevertheless is detectable in deep-sea deposits where the rate of sedimentation is very low. The surface receives energy from outside the earth in the form of solar radiation, nearly all of which is, however, reradiated into space.
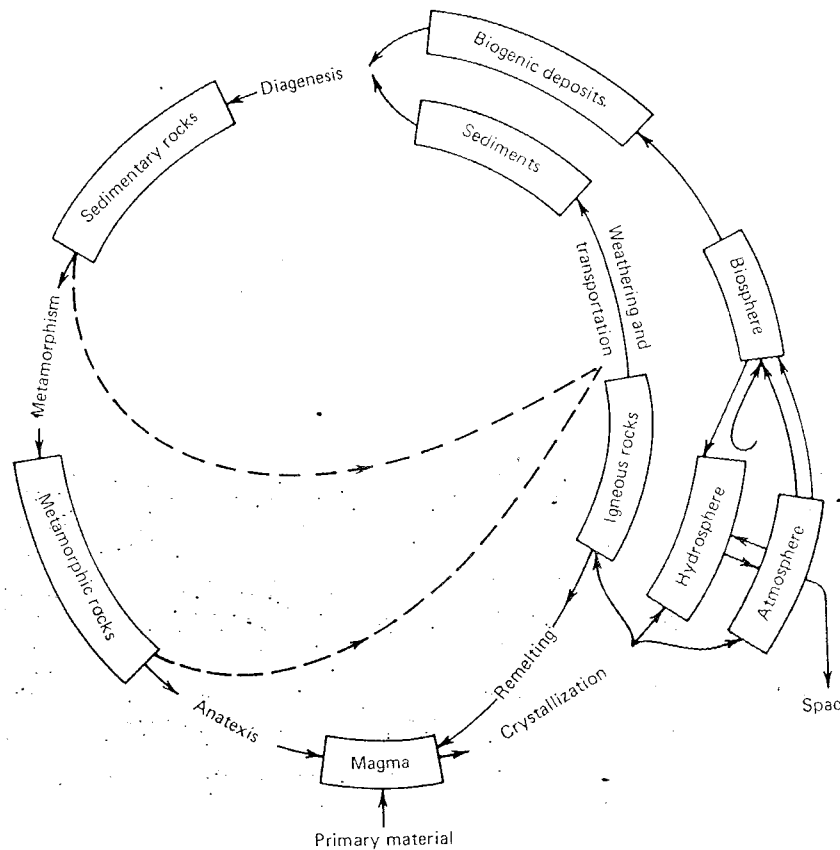
Figure 12.4 · The geochemical cycle.

Long-term changes in either the chemistry or mass of a unit or component in the geochemical cycle are known as secular changes; these include changes in the chemistry, mass, or isotopic composition of the continents as they evolve during geologic times. A major quest of geochemistry is to determine whether the continental masses are increasing in mass and if so how. One suggestion has been that andesitic magmas formed deep in the mantle along subducting plates at continental margins are added to the continental masses. Such andesitic rocks with intermediate rock compositions approximate the overall composition of the upper continental crust.

A consideration of the geochemical cycle of individual elements or groups of elements indicates that three main cycle types may be possible for an element at the earth's surface. In an ideal cycle neither the mass of a component nor its composition will change over geologic time. The cycle will be in a steady state with rates of influx and efflux matching. If material of different chemical composition is added to or removed from a cycle, it is possible that the cycle cannot return to its initial condition. Examples of such changes would be the slow addition of new magmatic material from the mantle to the crust or the addition of oxygen to the atmosphere. Material could also be added to or subtracted from the cycle without changing the overall chemical composition of the migrating material but changing its mass. Variations in the biomass with time could be an example of such a change. If a chemical change takes place, the cycle is said to have a secular chemical component, whereas if only a mass change takes place, the cycle has a secular nonchemical component.

The geochemical cycle provides a useful concept as a basis for the discussion of many aspects of geochemistry, particularly the course followed by a specific element in proceeding through the different stages. A complete understanding of the cycle in terms of the individual elements is one of the main objectives of geochemistry. An element may tend to concentrate in a certain type of deposit at some stage, or it may remain dispersed throughout the entire cycle. The geochemical cycle of a few elements has been worked out in considerable detail, but for many our knowledge is fragmentary.

In developing a geochemical cycle for a specific element it is important to remember that to balance the cycle not only are concentrations of the element in an individual component required but the rate of change or residence time is of critical interest.

THE HUMAN INFLUENCE ON THE GEOCHEMICAL CYCLE

In the course of human activities, particularly since the industrial revolution, perturbations in the natural geochemical cycles of some elements have developed. Some of the additions are simply greater contributions of compounds that have been in the natural geochemical cycles, whereas some are new materials produced by industrial processes. Abnormal fluxes of materials that cannot be accommodated by the geologic environment or produce undesirable effects in the biosphere are known as pollutants. Well-known pollutants include organo-mercury compounds, derived from industrial products such as fungicides in rivers, lead from gasoline, smog in urban areas, sulfur oxides near smelters, and agricultural nitrates and insecticides. Figure

Figure 12.5 · Factors involved in the flow of materials. Source: Robert U. Ayres and Allen V. Kneese, "Pollution and Environmental Quality," *Quality of the Urban Environment*, Harvey S. Perloff, Ed., Resources for the Future, Inc., Washington, D.C., p. 37.

Table 12.1 Primary Pollutant Sources and Amounts (Millions of tons/year), 1968

Pollutant Source	Weight of Pollutant Produced					Total Weight of Pollutant Produced by Each Source
	CO	NO$_x$	HC	SO$_x$	Part.	
Transportation	63.8	8.1	16.6	0.8	1.2	90.5
Fuel combustion (stationary sources)	1.9	10.0	0.7	24.4	8.9	45.9
Industrial processes	9.7	0.2	4.6	7.3	7.5	29.3
Solid waste disposal	7.8	0.6	1.6	0.1	1.1	11.2
Miscellaneous	16.9	1.7	8.5	0.6	9.6	37.3
Total weight of each pollutant produced	100.1	20.6	32.0	33.2	28.3	214.2

Adapted from U.S. Dept. of Health, Education, and Welfare, *Nationwide Inventory of Air Pollutant Emissions—1968*, p. 3.

12.5 illustrates schematically the complex path of materials and energy through an urban environment. Equally important are sources of agricultural pollution, and pollution due to mining activity.

Other than direct poisoning of water systems by contaminants such as mercury or lead, most pollution problems are due to the slow addition of materials to the atmosphere or hydrosphere. Several of these additions, such as the increased CO_2 contribution to the atmosphere, have been mentioned in earlier chapters.

Five types of primary pollutants account for over 90% of air pollution. The sources of these pollutants—carbon monoxide, nitrogen oxides, hydrocarbons, sulfur oxides, and particulates—are given in Table 12.1. Evaluation of these pollutants in terms of relative weighting factor or toxicity indicates that particulates are more dangerous than CO by a factor of about 100. Relative toxicities are estimated to increase in the order CO, hydrocarbons, sulfur oxides, nitrogen oxides, particulates. Water pollutants of geochemical interest include oxygen-demanding wastes, plant nutrients, organic compounds, inorganic chemicals including radioactive materials, sediments, and heat energy.

The addition of oxygen-demanding wastes can change the redox conditions of natural water systems and cause anaerobic microorganisms to replace aerobic microorganisms. An immediate result of such a change will be the production of H_2S in these waters instead of sulfate sulfur. The addition of excess plant nutrients, thereby increasing a lake's fertility, may produce the same effect. Eutrophication of a lake takes place by the accumulation of aquatic growth to the point where a lake is clogged with vegetation, or decaying vegetation uses up the available oxygen faster than it can be replenished. A commonly reported measure of the concentration of such wastes in water is the biological oxygen demand or BOD. In a water sample the BOD indicates the amount of dissolved oxygen required for the oxidation of these oxygen-demanding wastes. A BOD of 1 ppm is characteristic of normally pure water. Local pollution by industrial processes, especially the food industry, may increase the BOD as high as 10,000 ppm.

Inorganic water pollutants may include inorganic salts and mineral acids produced

by natural processes, as well as mining and manufacturing activities. Of particular geochemical interest is acid mine drainage. Mining activity that exposes sulfide-bearing rock to the atmosphere allows weathering to add significant concentrations of sulfuric acid and soluble compounds of iron to streams, especially in coal-mining areas. These substances are produced by the oxidation of FeS_2, in part perhaps with the interaction of bacteria.

Interesting studies of modeling of geochemical cycles have been made by Garrels and coworkers. Their model of the geochemical cycle of phosphorus shows the effect
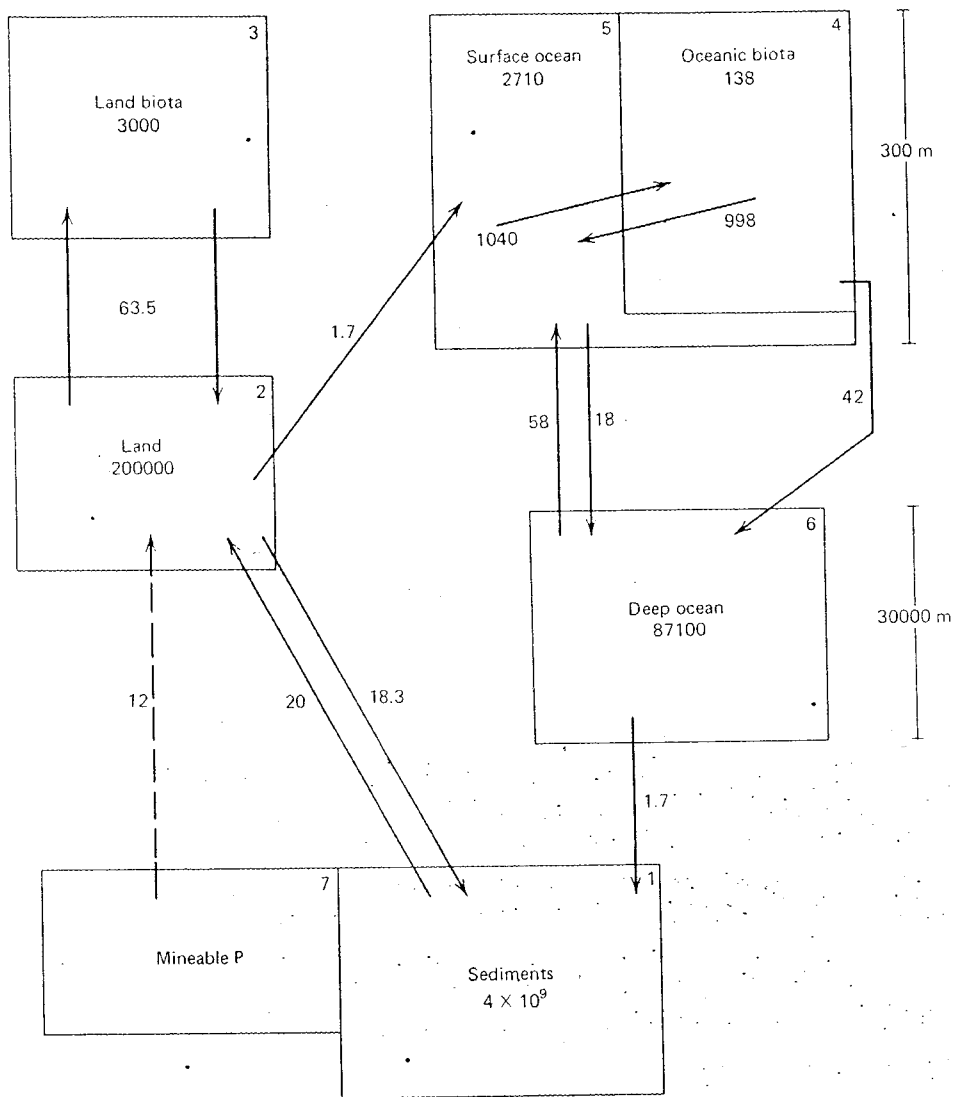
Figure 12.6 Global cycle of phosphorus (fluxes and reservoir masses in millions of metric tons P-units of 10^{12} g).

of real and theoretical changes on the cycle. Phosphorus is of interest because of its use as an inorganic fertilizer and its role in the eutrophication of waters. Addition to the geochemical cycle may be estimated by the records of the industrial production of phosphates, and since phosphorus does not enter the atmosphere via volatile compounds, the cycle is easier to model than nitrogen or carbon. Figure 12.6 shows a 1975 estimate of the reservoirs and fluxes of phosphorus. The selection of reservoir and flux values is complicated and may vary with individual studies, but it is essential for modeling studies that the flux rates in the cycle are balanced. The addition from the minable phosphorus is not included in the balance. In order to evaluate a perturbation in the cycle it is important to know that a flux or rate of reaction is directly proportional to the size of a reservoir if the reaction follows first-order reaction kinetics. Thus the rate constant for the flux from the ocean surface to oceanic biota will be 0.384×10^{12} g/yr and the return flux 7.232×10^{12} g/yr. If, as an example, the rate of photosynthesis were to decrease by 10% by a reduction in incident solar energy, the readjusted fluxes and reservoir masses could be calculated by utilizing rate constants and the fixed total phosphorus in the total surface reservoir. The new steady-state mass of phosphorus in the ocean surface would be approximately 2890×10^{12} g and the ocean biota 134×10^{12} g. The photosynthesis flux would only drop to 1011×10^{12} g P/yr due to the increase in ocean surface mass. The massive deep ocean reservoir and the upwelling of phosphorus to the ocean surface act to buffer the system against rapid change. In systems with smaller reservoirs such as a lake, natural or artificial perturbations may cause relatively greater changes in the natural cycles of elements.

ENERGY CHANGES IN THE GEOCHEMICAL CYCLE

We have been considering the geochemical cycle in terms of the material changes that take place during the various processes. Equally significant, if less studied and less well understood, are the energy changes during the cycle. Geochemical processes operate only because of a flow of energy from a higher to a lower potential or intensity; hence energy is no less important than matter in the geochemical cycle.

The earth cannot be considered a closed system in terms of energy, since it receives a large amount of heat from the sun in the form of solar radiation. In an overall view we should consider the whole universe as a single system that is presumably undergoing a spontaneous decrease in free energy and an increase in entropy, as prescribed by the second law of thermodynamics. However, we can deal with the principal aspects of the geochemical cycle in terms of the solar system, since the forces acting on the earth and the energy it receives are primarily the result of its relationship to the sun. Within this limited system the overall development is also toward a decrease in free energy and an increase in entropy. The few apparent exceptions, such as the accumulation of free energy in deposits of coal and oil, represent a decrease in entropy at the expense of the sun and are merely minor fluctuations in a general trend toward a state of higher entropy.

Major sources of energy in the geochemical cycle may be considered under the following headings: (a) solar radiation; (b) mechanical energy, potential and kinetic;

(c) reaction energy; (d) nuclear energy; and (e) heat content of the earth. Transformation of these forms of energy is the common denominator of all geochemical processes and reactions.

Solar radiation is, of course, the product of nuclear reactions whereby hydrogen is converted to helium in the sun. The sun radiates energy in all directions and only a very small proportion, about 5×10^{-8}, impinges on the earth; about half of this is reflected back into space by the atmosphere. The average solar radiation received at the earth's surface is 4.2×10^{-3} cal/(cm^2)(sec), or about 6750×10^{20} cal annually for the whole earth. It has been estimated that 0.1% is utilized by photosynthesis of plants, terrestrial and marine, about 15% is absorbed by the earth, and the remainder is used to evaporate water from the hydrosphere. Solar energy is thus essential to the cycle of changes involving the interaction of the hydrosphere, the biosphere, the atmosphere, and the crust.

Through the evaporation of water from the hydrosphere solar radiation is converted into mechanical energy. Water vapor in the atmosphere has potential energy, which in turn is transformed into the kinetic energy of rain and running water, the source of most of the energy driving erosion and sedimentation. This kinetic energy is utilized in part in moving solid material downhill. A pound of quartz moved from the top of Pikes Peak to the delta of the Mississippi River loses about 14,000 foot-pounds of energy; its original potential energy is converted into kinetic energy and eventually into heat. The reduction of a mountain range to a peneplain is largely the work of running water and transforms the energy stored in its uplift into heat. Within a comparatively short period of time, geologically speaking, erosion would reduce the whole land surface to a monotonous low-lying plain, were it not for diastrophism, which creates mountains about as fast as they are destroyed. The uplift requires a considerable expenditure of energy. The average elevation of the continents is about 800 meters; to maintain this average elevation against erosion requires an output of 10^{17}–10^{18} cal/year.

Another source of mechanical energy is the kinetic energy of the earth's rotation. A portion of this energy appears as winds and tides. To be sure, tides are brought about by the gravitational effect of the moon and, to a lesser extent, the sun; but the energy dissipated comes from the kinetic energy of the earth and not from these extraterrestrial bodies. The effect of these tides is the slowing-down of the earth's rotation by about one second in 1000 years. The kinetic energy of the earth's rotation has important geological implications, for it acts to maintain the earth's figure against all forces that tend to deform it.

Reaction energy covers all changes of state, physical and chemical, and is the form most immediately associated with the geochemical cycle, because every geochemical process involves such changes. From Le Châtelier's principle we can predict that reactions taking place under conditions of falling temperature, such as the crystallization of magma, will be exothermic; reactions taking place under conditions of increasing temperature, such as progressive metamorphism, will be endothermic. Reactions involving change of state, such as the melting of solid material and polymorphic transformations, are accompanied by important energy changes, and are probably significant in the overall energy balance within the crust and the mantle.

Table 12.2 Radioactive Heat Production in Typical Rocks

	Concentration			Heat Production	
	U (ppm)	Th (ppm)	K (percent)	cal/(g)(10^6 yr)	($t = -4.5 \times 10^9$ yr)
Granitic	4	16	3.5	8	33
Intermediate	1.7	7	2	3.5	14
Basaltic	0.5	2	0.9	1	5
Peridotite/ dunite	0.01	0.04	0.001	0.01	0.04

We have already seen that solar radiation is the product of nuclear reactions in the sun. Nuclear reactions in the earth—the spontaneous disintegration of radioactive atoms—have been of prime importance throughout the earth's history. Radioactive decay converts mass to energy, and nearly all the energy appears as heat in the immediate neighborhood of the source. The isotopes of importance for heat generation in the earth are ^{238}U, ^{235}U, ^{232}Th, and ^{40}K. Table 12.2 gives some typical figures for radioactive elements in selected rock types, together with calculated rates of heat production. These figures are based on averages of numerous analyses; however, the individual analyses show a considerable variation, and the sampling problem when dealing with uranium and thorium, present in a few parts per million, associated with minor mineral constituents and irregularly distributed, is a formidable one. Nevertheless, the results show quite clearly the marked concentration of the radioactive elements in the silica-rich rocks of the continental crust.

The figures in Table 12.2 are based on the amounts of uranium, thorium, and potassium in the rocks at the present time. The present amounts of the radioactive isotopes are smaller than those that existed in the past, and for calculations dealing with the whole lifetime of the earth it is important to take the decay into account. The rate of heat production at a time t years ago was $e^{\lambda t}$ times the present rate, for a given present quantity of an isotope with disintegration constant λ. The ratio of the total heat produced over a time t, allowing for the decay, to the amount that would have been produced in the same time if the present quantity had been constant, is $(e^{\lambda t} - 1)\lambda t$.

Table 12.2 includes the calculated total heat generation for each rock type 4.5 $\times 10^9$ years ago. These figures point up the significantly greater production of radiogenic heat in the early period of the earth's history. A major proportion of the early heat was contributed by ^{40}K.

The heat content of the earth is due to the thermal vibration of its constituent atoms. The internal heat of the earth is enormous, but much of it is essentially static and plays no part in the geochemical cycle at and near the surface. Thermal energy is transferred to the surface by conduction and by igneous activity; much of this energy is generated within the crust by radioactivity. The continental crust averages about 35 km thick. A 35 km column of granite with 1 cm^2 cross section has a volume of 3.5 $\times 10^4$ cm^2 and a mass of 9.4 $\times 10^6$ g. The heat generation, from Table 12.1,

would thus amount to 75 cal/(cm^2)(year), or approximately 2×10^{-6} cal/(cm^2)(sec). The average output of heat by conduction is about 1.4×10^{-7} cal/(cm^2)(sec), or roughly 3×10^{20} cal/year for the whole earth. A granitic crust 35 km thick would thus supply more heat than is actually escaping at the surface, even without allowance either for original heat or for radiogenic heat liberated in the underlying mantle and core.

Precise estimates of the heat brought to the surface by igneous activity are difficult to make. Verhoogen (1960) estimated that about 3×10^{15} g of lava are erupted annually, giving off some 400 cal/g during cooling and solidification, that is, a total of about 10^{18} cal. Not all magma reaches the surface; erosion reveals igneous rocks of all ages that have been intruded in the liquid state and have cooled within the crust. In addition, volcanic heat is also carried to the surface by thermal springs. The total amount of heat contributed by igneous activity is thus of the order of 10^{18}–10^{19} cal/year. Although the contribution of a single volcano may be spectacular, the total heat transferred by igneous activity is only a small fraction of that due to conduction. The average energy released annually by earthquakes has been estimated to be about 10^{25} ergs, or 0.24×10^{18} cal, also a minute fraction of the heat escaping by conduction.

In terms of the geochemical cycle the upward movement of magma brings with it a certain amount of energy. Our knowledge of the thermal gradient within the earth indicates that temperatures within the crust and mantle are on the average well below the temperature of melting of the material at any specified depth. The thermal energy of the earth (including that produced by radioactivity) provides an ample source of heat; the problem is to find some mechanism that will cause melting at local "hot spots," the centers of igneous activity. Igneous activity is a small-scale, local phenomenon in relation to the earth as a whole, and it is concentrated along zones of general tectonic activity, especially the margin of the Pacific Ocean and along midocean ridges. In many places igneous activity is related to deep-seated fractures; these fractures may act as channels for gases carrying heat from greater depths, and movements on such fractures will disturb the balance between temperature and pressure on either side. Such mechanisms may produce magma, which by virtue of its lower density than the surrounding solid material tends to rise into the higher and cooler parts of the crust. The thermal energy of the magma is dissipated to the surroundings during cooling and solidification. The kinetic energy of the atoms or ions decreases as the temperature falls; their potential energy also diminishes when they are arranged in an orderly fashion in a crystal lattice. At the next stage of the cycle, erosion and sedimentation reduce the material on which they work to a lower energy state. Erosion moves material with the aid of gravitational forces, and a loss of potential energy results. The chemical reactions accompanying weathering are spontaneous and usually irreversible and lead to a decrease in free energy. The only spontaneous process at the earth's surface resulting in an increase of free energy is the conversion of carbon dioxide and water to complex organic compounds by plants utilizing solar radiation. The general trend is reversed during metamorphism, when thermal and gravitational energy is converted into chemical energy by endothermic reactions and the formation of compounds of higher density. A sufficient increase in

temperature may increase the kinetic energy of the atoms or ions sufficiently to overcome the forces holding them in crystal lattices, so that the minerals decompose or melt. A magma is thereby regenerated, and the cycle is complete. A quantitative balance sheet for the energy change during the geochemical cycle cannot yet be drawn up, no more than a quantitative balance sheet for the material changes. However, it does appear that much of the energy of the geochemical cycle is contributed by the disintegration of radioactive elements within the crust, possibly supplemented by some of the earth's internal heat.

SELECTED REFERENCES

Engel, A. E. J., S. P. Itson, C. G. Engel, D. M. Stickney, and E. J. Cray (1974). Crustal evolution and global tectonics: a petrogenic view. *Bull. Geol. Soc. Amer.*, 85, 843–858.

Garrels, Robert M., F. T. Mackenzie, and C. Hunt (1975). *Chemical cycles and the global environment*. Wm. Kaufman, Los Altos, Ca. An assessment of human influences on exogenic geochemical cycles from a geochemical viewpoint.

Geochemistry and the environment. Vol. 1 (1972) 113 pp.; Vol. 2 (1977) 163 pp.; Vol. 3 (1978) 200 pp. National Academy of Sciences, Washington. The relation of selected trace elements to health and disease.

Goldschmidt, V. M. (1954). *Geochemistry*. Part II of this book deals with geochemical cycle of individual elements and groups of elements.

Jakes, P. (1973). Geochemistry of continental growth. In *Implications of continental drift to the earth sciences*, D. H. Tarling and S. K. Runcorn (eds.), Vol. 2, pp. 991–1001. Academic Press, New York.

Mueller, R. F. and S. K. Saxena (1977). *Chemical petrology* 394 pp. Springer-Verlag, New York. An advanced text that includes applications of chemical petrology to the terrestrial planets and meteorites.

Rankama, K. and Th. G. Sahama (1950). *Geochemistry*. In part II of this book the geochemistry of the individual elements is described, including what is known of the geochemical cycle for each of them.

Shaw, D. M. (ed.) (1963). Studies in analytical geochemistry. *Roy. Soc. Canada Spec. Publ. 6.* 139 pp. A number of review articles, including one on sulfur isotope geochemistry by H. G. Thode and one on oxygen isotope geochemistry by R. N. Clayton.

Stoker, H. S. and S. L. Seager, (1976). *Environmental chemistry* (second ed.). Scott, Foresman and Co., Glenview, Ill. A good source of data on air and water pollution.

Wedepohl, K. H. (ed.) (1969). *Handbook of Geochemistry* Springer-Verlag, Berlin. An important compilation of the geochemical data on each element. This multi-volume loose-leaf book is design-d to be updated by replacement and addition of sections.

Wyllie, P. J. (1971). *The dynamic earth*. 416 pp. John Wiley & Sons, New York. An advanced overview of the solid earth including geophysical, tectonic, and petrologic considerations.

APPENDIX 1

Radioactive Decay Schemes for Uranium and Thorium

Element	U-238 Series						
Neptunium							
Uranium	U-238 4.468 × 10^9 yrs		U-234 2.45 × 10^5 yrs				
Protactinium		Pa-234 1.175 min					
Thorium	Th-234 24.10 days		Th-230 8.05 × 10^4 yrs				
Actinium							
Radium			Ra-226 1,622.0 yrs				
Francium							
Radon			Rn-222 3.8235 days				
Astatine							
Polonium			Po-218 3.05 min		Po-214 164 μ sec		Po-210 138.38 days
Bismuth				Bi-214 19.7 min		Bi-210 5.01 days	
Lead			Pb-214 26.8 min		Pb-210 22.3 yrs		Pb-206 stable lead (isotope)
Thallium							

99.98%

	Th-232 Series				U-235 Series			

Th-232 Series / U-235 Series decay chart

Th-232 Series					U-235 Series			
					U-235 7038 X 10⁹ yrs			
						Pa-231 3.43 X 10⁴ yrs		
Th-232 1.41 X 10¹⁰ yrs		Th-228 1.9131 yrs			Th-231 25.52 hrs		Th-227 18.718 days	
	Ac-228 6.13 hrs					Ac-227 22 0 yrs		
Ra-228 6.7 yrs		Ra-224 3.66 days					Ra-223 11.435 days	
		Rn-220 55.6 sec					Rn-219 3.96 sec	
		Po-216 .15 sec	65%	Po-212 0.29 μsec			Po-215 1.78 X 10⁻³ sec	
			Bi-212 60.48 min					Bi-211 2.15 min
		Pb-212 10.64 hrs	35%	Pb-208 stable lead (isotope)			Pb-211 36.1 min	Pb-207 stable lead (isotope)
			Tl-208 3.053 min					Tl-207 4.77 min

INDEX