The Earth and Planetary Sciences

George W. Wetherill and Charles L. Drake

Summary. During the last two decades the earth sciences community has become persuaded that the earth is a dynamic body; an engine driven by its internal heat. The major surface manifestation of this dynamism has been fragmentation of the earth's outer shell and subsequent relative horizontal movement of the pieces on a large scale. The driving force is convection within the earth, but much remains to be learned about the nature of the convection and the composition of the earth's interior. The other terrestrial planets show evidence of once having been hot, but their surfaces suggest long-term stability and lack evidence of continuing convection.

The Earth as a Machine

The earth is many things to many people, and is viewed in different ways by farmers, explorers, prospectors, poets, and scientists. Probably its most significant image is that of the only known abode of intelligent life, whereas from the physical and chemical point of view it is probably not such an unusual planet. Yet the physical planet represents the foundation upon which life has developed, and understanding its physical and chemical nature is of fundamental importance not only to geologists but to biologists and other scientists as well.

Looked at from the purely physical point of view, the earth has often been loosely described as a "heat engine." It is now clear that this is not simply a metaphor but that it is literally true. Much of the present research in earth science can be thought of as an attempt to understand how this complex thermomechanical system operates. Moreover, this approach leads naturally to consideration of the things we can learn about the earth by studying other planets.

The moving parts of the engine are the material of the earth itself: the lithosphere, consisting of the solid crust and outermost mantle; the relatively plastic deeper interior; and the fluid atmosphere, seas, and outer core. The energy sources that drive the engine are principally nuclear: the thermonuclear solar radiation and the radioactive disintegration of long-lived isotopes in the earth's interior, augmented by relics of the gravitational energy released when the earth was formed 4.5 billion years ago. Quantitatively, the input of solar energy to the earth is much greater than that generated internally, about 10^{12} megawatts compared with 10^7 MW. The movement of wind and water, and therefore many surface geological processes, depends on this solar energy. However, this source does little to heat the earth's interior. Even at the relatively low temperature of the earth's surface, the thermal loss from blackbody radiation is sufficient to very nearly balance the solar input. As far as the internal processes of the earth are concerned, the ~ 300 K equilibrium surface temperature simply serves as a relatively cool, surface boundary condition.

The more feeble internal energy generation is of much more significance to the internal dynamics of the earth. The heat remains trapped within the rocks in which it is generated until the accumulation of energy drives the temperature near to or above the melting point. At these elevated temperatures the viscosity of the material becomes sufficiently low to permit convective transport of heat to the surface. It must be admitted, that when viewed as a machine, the earth we value so much is a very inefficient one. Although the rate of energy production in the earth's interior by radioactive decay is comparable to the rate of energy consumption by human civilization, the kinetic energy associated with the resulting movement of the great plates of lithosphere is less than that required to heat the water for a cup of tea. Nevertheless, over the many hundreds of millions of years of the earth's history, this gradual motion has rearranged the continents and ocean basins so much that it is difficult to reconstruct with much certainty the geography of the earth's surface further back than a few percent of the age of the earth.

Until this century, geological study was effectively limited to surficial portions of the earth's continents—to the interface between the region dominated by processes driven by the flux of solar energy and those resulting from internal movement. The nonlinear interaction of the great variety of resulting phenomena has entangled the effects of these two regimes, and their untangling has been, and continues to be, a major challenge to the skills of earth scientists. For example, prior to not much more than a decade ago it was entirely respectable to maintain that mountain building was primarily driven by surface processes, that is, it was the complex end result of the surface processes of deposition of sediments in basins bordering the continents, such as the Gulf of Mexico. A more up-to-date example is the fact that we still remain unable to distinguish with any great degree of confidence the geochemical signature of a mantle source of continental igneous rocks from that superimposed by interaction with the more near-surface environment.

The Plate Tectonic Revolution

During the last 20 years there has been a revolution in the science of the earth. This revolution grew from minority views that, through the force of evidence and persuasion, became those of the majority. While emphasis in the past had been on the forces and the vertical movements that had produced the topography and surface structure of the globe, it now appeared that these were the relatively minor by-products of major horizontal movements of very large fragments of the earth's outer shell. As these fragments, or plates, moved relative to each other, they produced mountain systems where they collided, rifts and new crust where they separated, and major strike-slip faults, such as the San Andreas fault of California, where they slid relative to each other (Fig. 1).

The basic concept of the relative movement of continents was not new; what was new was the fresh evidence that led to its acceptance. Much of this evidence came from studies of geomagnetism. It has long been known that the

Dr. Wetherill is director of the Department of Terrestrial Magnetism, Carnegie Institution of Washington, Washington, D.C. 20015. Dr. Drake is professor of earth science at Dartmouth College, Hanover, New Hampshire 03755.

 0036-8075/80/0704-0096$02.00/0

earth's magnetic field, rather than originating from a permanent magnet, has a dynamic source, because it varies in strength and position. The most logical source is the earth's core, since seismology tells us that the core behaves as a liquid and is made mostly of metallic iron, a conductor. Fluid motions in a conducting medium of this sort in a weak preexisting field are capable of amplifying this field to produce the observed field. Although the symmetry of the motion requires the field to be roughly oriented along the axis of rotation, the polarity of the field is another matter. There is no good reason why it should be in the present direction, rather than in the opposite direction.

During the early 1960's careful measurements of the polarity of the magnetic field preserved in rocks, together with the use of new methods for dating the time at which they became magnetized, demonstrated that the magnetic field reversed itself on a time scale of 10^5 to 10^6 years. Measurement of the reversal times in rocks on different continents showed that these reversals were simultaneous phenomena on a global scale. This permitted establishment of a "reversal time scale" for the rocks of the earth. Thus it was found that magnetic field variations at sea, measured perpendicularly to the axis of the mid-ocean ridge system, repeated the same patterns in space that the reversal sequence did in time. Thus, it was concluded the sea floor is a giant magnetic tape recorder and that new crust becomes magnetized with the polarity of the field at the time of its formation. Together with the solid outermost mantle, the crust then moves away from the ridge axis at a rate of a few centimeters per year. This is consistent with the observation that most of the present-day earthquake and volcanic activity within the ocean basins takes place in association with the ridge axes where new crust is being created.

During the last decade the Deep Sea Drilling Project and its successor, the International Phase of Ocean Drilling, have had most remarkable success in defining the history and character of the

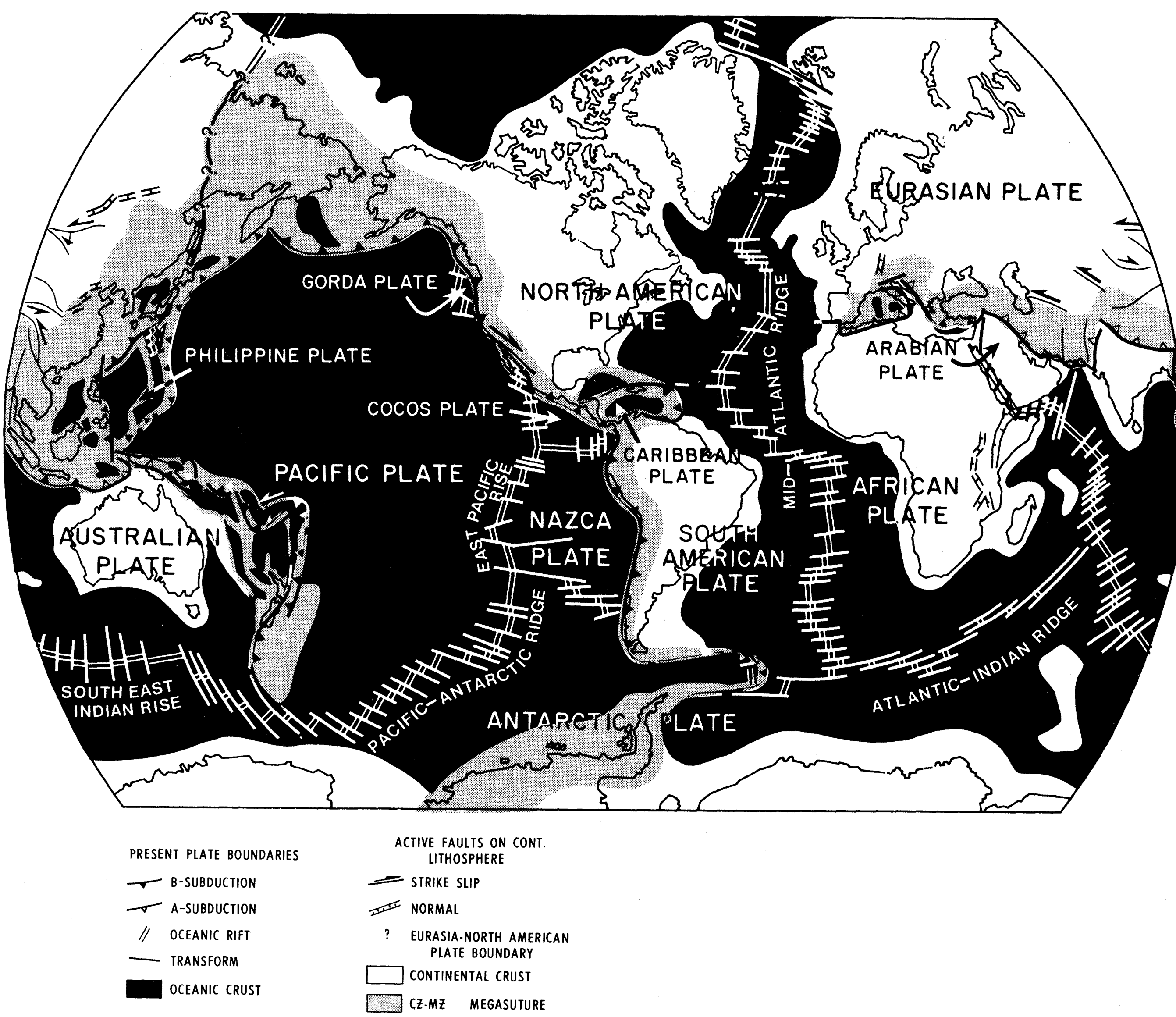

Fig. 1. Lithospheric plates. The oceanic areas contain crustal rocks formed at mid-ocean ridges during the last 200 million years. Corresponding crust created or deformed by compressional processes during the same period is contained in the Cenozoic-Mesozoic (*CZ-MZ*) megasutures. These megasutures are usually bounded on one side by the dipping zones of earthquakes associated with the underthrusting of oceanic lithosphere (*B-subduction*) and on the other by underthrusting unaccompanied by major earthquake activity on the continents (*A-subduction*). They contain the young mountain systems of the world. Current volcanic and earthquake activity is concentrated along ocean ridges or along these megasutures. These active zones divide the outer shell of the earth, the lithosphere, into a number of spherical plates moving relative to each other. The sum of these movements is plate tectonics. [From A. W. Bally, Shell Development Company]

ocean crust, the nature and distribution of the sediments on the sea floor, and the sedimentological, climatological, and oceanographic consequences of the plate tectonics model (Fig. 2). One of the first contributions of these endeavors was to confirm the inference from the magnetic data that the age of the ocean crust increases with distance from the ridge axis.

To accommodate the new material injected along the ocean ridges without expansion of the earth, old lithosphere must be eliminated (Fig. 3). This process takes place in the other major earthquake and volcanic belt of the earth, that rimming the Pacific Ocean and extending from Indonesia through the Himalayas to the Alps. In this belt earthquakes occur to depths in excess of 700 kilometers, and seismological evidence indicates that in these regions oceanic crust and lithosphere are underthrust to similar or greater depths. Rates of motion are more difficult to establish in these areas, but if they are of the order of a few centimeters per year, then older material can be disposed of at rates compatible with the rates of generation at the ridge axis.

The plate tectonics model very effectively explains the history, the nature, and even the gross topography of the oceanic crust that covers about 70 percent of the earth's surface. This thin (about 5 km thick) oceanic crust overlies the solid oceanic lithosphere, which in turn lies above the more plastic and partially melted asthenosphere. Sea floor topography can be explained by a relatively simple thermal model that is compatible with seismological data on the thickness of the lithosphere and data on the flow of heat through the sea floor. As the crust and lithosphere move away from the ridge, the lithosphere cools, and this cooling converts asthenosphere into lithosphere. The resulting thicker and denser layer of crust and upper mantle causes the elevation to decrease as the age of the ocean crust increases with distances from the ridge. The thickening is rapid at first and asymptotically approaches a lithosphere thickness of about 85 km. Thus the ocean ridges are not discontinuous masses superimposed on the basins, but represent those youngest parts of the oceans in which the elevation is most apparent.

This new interpretation of oceanic topography has implications that extend to the continents. It implies that the area of elevated sea floor adjacent to the ridge may be a function of the rate at which new oceanic crust is generated. Since this rate is not constant, we should expect, during periods of rapid generation, seawater to be displaced and to flood the land, and the continents to appear as many islands surrounded by shallow seas. During periods of slow crustal generation, the seas should retreat from the continental areas. This interpretation may explain the last major flooding of the continents, which occurred 80 million to 100 million years ago and has been related to a period when rapid generation of new crust is believed to have occurred.

Fig. 2. The Hughes *Glomar Explorer*. This vessel is under consideration as a replacement for the *Glomar Challenger* which has been drilling in the deep sea since 1968. The large size of *Explorer* would permit the use of riser pipe and blowout preventers to control the hole. As a result, *Explorer* would be able to drill the thick sediments of the deeper parts of the continental margins. These regions are of great interest scientifically. The passive margins hold the history of ocean basin formation; the active margins contain the record of deformation associated with converging lithospheric plates. *Explorer* also has the potential for drilling deeper into the ocean crust than was possible from *Challenger* and of obtaining paleoceanographic data from high latitudes. [Courtesy of Robert Bauer, Global Marine]

Formation of the Oceanic and Continental Crusts

Laboratory studies of the mineralogy and chemical composition of these ocean ridge basaltic volcanic rocks show that they are remarkably uniform, even on a global scale, and are thus derived from a relatively homogeneous, partially melted, underlying asthenospheric mantle. The trace element chemistry and isotopic composition of the basalts show that their mantle source is not characteristic of the earth as a whole, but has been greatly depleted in the "large ion lithophile elements," potassium, rubidium, strontium, barium, uranium, thorium, and rare earth elements. This depletion appears to have been taking place for much longer than the 200 million years of age of the present ocean basins, for at least as long as 1700 million years, and probably much longer. In contrast, basaltic rocks from oceanic islands such as the Azores and Samoa cannot be obtained in their entirety from this same source. They bear the signature of a source that is more heterogeneous and less depleted in basalt-forming large ion elements than the source of the more abundant ridge basalts. The reasons for these differences are poorly understood. It is commonly supposed that the sources of the oceanic island basalts are deeper in the mantle, and that partially melted material rises to the surface in "plumes" or "blobs," passing through the well-mixed source region of the ridge basalts.

The oceanic areas are young, representing only 200 million years, the last 5 percent of earth history. Therefore they are thermally immature and relatively simple to model since they have been affected by a limited number of tectonic events and processes. Because of their lower mean density, the continental regions are more buoyant. When a lithospheric plate surmounted by continental crust collides with oceanic lithosphere, it is the denser oceanic lithosphere that is subducted. Thus continents can survive over major portions of earth history; ocean basins have only a transient existence.

The continents are therefore much more complex. They contain what remains of the first 95 percent of the earth's history. Most of continental areas have been affected by multiple tectonic events, and continental surfaces have been continuously subjected to the weathering and erosional processes driven by the more intense input of solar energy. These surficial geological processes have low-temperature chemical effects comparable in magnitude and more complex in nature than the chemical fractionation that accompanies the formation of oceanic basalts. It is difficult to assess the extent to which the granite-rich nature of the continental crust is a direct consequence of introduction of material from the mantle, and the extent to which surficial chemical processes are also involved: those accompanying sedimentation, metamorphism, and remelting of sediments and more deep-seated crustal rocks to form granitic rocks. In spite of such complexities these rocks of the continents must be understood if we are ever to understand more than the tectonics of the relatively recent past. The old oceans have vanished forever, and knowledge of the ancient earth will now require deciphering the complex continental geological record. This record is not only complex, but incomplete. Much of the structural history of the continents is hidden beneath relatively undisturbed sediments (Fig. 4).

A great many holes have been drilled on the continents, but the great majority of these are either quite shallow, or are drilled in sedimentary basins in the search for hydrocarbons. In recent years, drilling programs have been initiated for a variety of purposes and it may be possible to use the holes drilled for resources as well as others drilled exclusively for scientific purposes to reduce our ignorance of the distribution, character, and history of the early continental crustal rocks. The scientific usefulness of such holes can be increased, and costs of drilling reduced, if adequate geophysical and geological studies are conducted both before and during drilling. We have learned from studies of deep seismic reflection, for example, that our projections of surface geology in the third dimension are often at variance with reality.

A fundamental half-truth of geology is the "uniformitarian" doctrine which holds that "the present is the key to the past." Insofar as this is valid, it provides a starting point in our attempt to understand the processes that formed the continents. By analogy we can observe the processes at work in modern zones of plate convergence by studying island arcs and the young mountain systems that are the products of these collisions between plates.

In such regions, volcanic rocks are being formed by partial melting of denser material more similar in chemical composition to the bulk earth than to the ridge basalts. These volcanic rocks have lower density than those formed at the

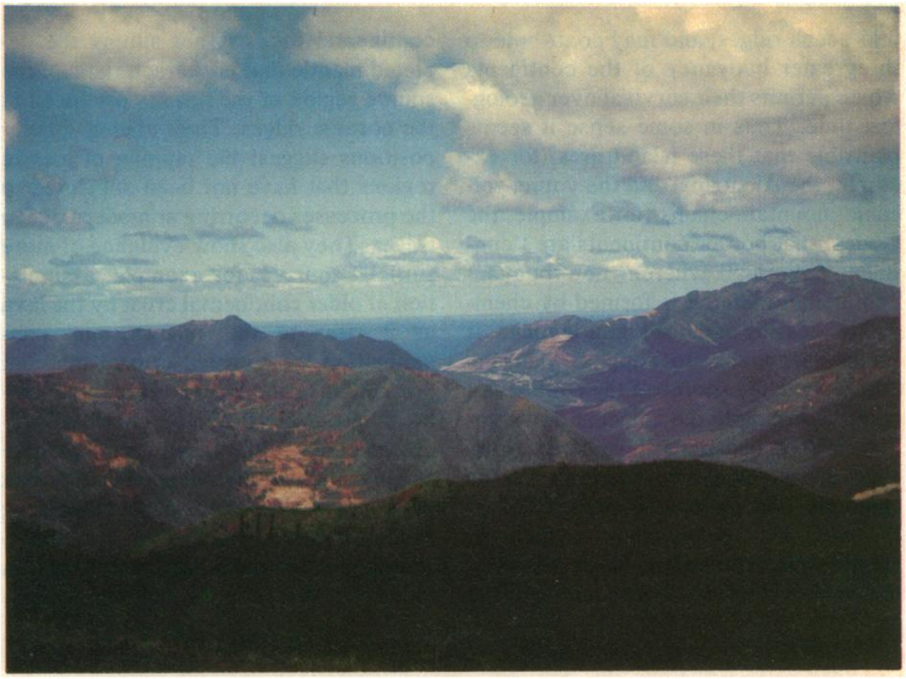

Fig. 3. New Caledonia. The elevated backbone of New Caledonia consists of rocks from the oceanic lithosphere. It represents a slice of lithosphere overthrust onto the island about 30 million years ago. This feature, produced by plate tectonics, is of some economic importance since the rocks of New Caledonia are a major source of nickel. [Photo by C. Drake]

Fig. 4. Mule Ear diatreme. On the edge of the Monument Uplift in the Colorado Plateau, Mule Ear diatreme is one of a number of windows into the earth's interior in the western United States. These and similar explosive vents in other parts of the world bring up rocks from as deep as several hundred kilometers and provide direct opportunities for sorting out the nature of the deep crust and upper mantle. [Photo by C. Drake]

mid-ocean ridges, and may contribute to the greater buoyancy of the continents which permits their survival over geological time. Thus in some sense it seems plausible that these island arcs (for example, the Aleutians) and the young volcanic mountain chains (for example, the Andes) that border continents are "continent factories" wherein new buoyant continental material is formed by chemical processing of denser mantle rocks. It seems reasonable to suppose that the partially melted source of the continental material is simply the adjacent down-going slab of oceanic lithosphere. However, geochemical investigations of trace elements and isotopic compositions of the volcanic rocks argue against so simple an interpretation. These measurements show that the source region of the continental rocks is not simply the depleted mantle that is characteristic of the source region of the basalts produced at the oceanic ridges. Their chemical compositions suggest the tapping of mantle regions that have not been subjected to the processes occurring at modern ocean ridges. They also show evidence of more surficial sources, for example, assimilation of older continental crust by the lava as it rises to the surface, and contamination by recycled crustal material present in subducted sediments. Untangling the physical and chemical processes occurring in these regions where oceanic and continental material interact with one another represents a continuing challenge at the frontier of our understanding of the earth, and will remain a focus of research in the coming decade.

Before the advent of modern plate tectonics, much emphasis was placed on the significance of the approximately 35-km-thick continental crust and its boundary with the underlying mantle, the Mohorovicic discontinuity. During the early years of the plate tectonic revolution, there was a tendency to downgrade the fundamental significance of continents, the view being that they were relatively thin pieces of superficial "scum" attached to the truly significant lithospheric plates. This view is now changing. Although our knowledge of the continental lithosphere is far from adequate, it seems as though the continent versus ocean distinction may be even more profound than believed earlier. The plates which underlie continental crust appear to be more slowly moving relative to one an-

Fig. 5. Meteor Crater, Arizona. Most of the terrestrial planets and the moon have heavily cratered surfaces resulting from a heavy bombardment of extraplanetary objects that diminished greatly in intensity about 3900 million years ago. Many fossil craters have been found in rocks up to 2000 million years in age on the earth. Among the younger of these is Meteor Crater in Arizona. The relative scarcity of craters on the earth as compared to the moon is a reflection of the dynamic nature of the earth. Since it is very unlikely that the earth was spared the bombardment, the record must have been destroyed by subsequent tectonic activity and erosion. Thus, in contrast with the other planets, the surface features of the earth are geologically relatively young. [Photo by D. Roddy, U.S. Geological Survey, Flagstaff, Arizona]

other and, at least in some regions, much thicker (200 to 400 km) than their oceanic counterparts. From the viewpoint of the geochemist, continental mantle sources give evidence of being much less differentiated relative to the bulk earth than is the case for the mantle source of ocean ridge basalts.

The Interior of the Earth

The horizontal movement of the lithospheric plates, both those bearing continents and those purely oceanic, and the subduction of the latter at plate boundaries, represent the most overt expression of the earth's internal heat engine. It cannot be all there is, since in order to conserve matter, there must be a more deep-seated counterflow, whereby matter is returned to the spreading centers at the ridges. The nature of this counterflow is obscure at present. In some sense the sea-floor spreading, the descending subducted slabs, and the counterflow must be portions of a system of convective flow. Like other convective systems, differential buoyancy is the driving force, and the earth's internal heat is the source of energy that powers the convection. The differential buoyancy results from ordinary thermal expansion as well as differences in density associated with phase changes, both between solid and liquid and those involving only solid-solid phase transitions. However, it has proved difficult to go much beyond broad generalizations of this sort. Very little is known about the counterflow. At least some of it must extend to depths of greater than 700 km because deep-focus earthquakes show that the descending slabs penetrate to depths at least that great. Dynamic studies of stress balances show that the principal counterflow cannot be confined to the weak asthenosphere immediately underlying the lithosphere. For the most part at least, both the lithosphere and asthenosphere belong to the same uppermost portion of the convective system.

Although it is clear that in a general sense the plate tectonics motion of the earth represents a convective system, it differs in many ways from more familiar convective systems studied in the laboratory, treated by classical theories, or observed in the atmosphere. For one thing, the convective motion cannot be described as a steady-state motion of symmetric convective cells. The particular pattern of sea-floor spreading observed at present cannot be older than the age of the present ocean basins, about 200 million years. This is the time required for a plate to move 10,000 km, one-fourth of the earth's circumference, at typical spreading velocities of 5 centimeters per year. Therefore, if this movement is thought of as a convective cell, there has not been enough time to complete one cycle of periodic motion. In addition, even if the counterflow extends into the lower mantle of the earth, the ratio of the horizontal to vertical dimensions of the earth's convection "cells" is markedly greater than the approximately equidimensional flows studied in the laboratory.

Some of the reasons for these differences are probably related to the great differences in pressures, density, temperature, and viscosity of the various portions of the convective system. A full understanding of the dynamics of the earth's internal heat engine must also incorporate what has been learned during the last 30 years about the chemistry and mineralogy of the earth's deep interior.

Seismologists have long known that the internal constitution of the earth could not be simply explained in terms of a homogeneous silicate mantle and core. Over a major portion of the mantle, between depths of 400 and 1000 km, it proved impossible to explain the observed velocities of elastic waves as entirely the consequence of compression of chemically and mineralogically homogeneous material. Even after proper allowance was made for the effects of pressure and temperature, an anomalous increase in velocity and inferred density was found. It was hypothesized that these effects were caused by solid-state phase transitions whereby the familiar forms of crustal silicate minerals were transformed into dense and therefore more thermodynamically stable minerals under the pressures of hundreds of thousands of atmospheres that occur at these depths. A beautiful sequence of laboratory and theoretical investigations has identified the most probable mineralogical nature of these phase changes as well as their relation to the "discontinuities" in densities and velocities identified by seismological work. At relatively shallow depths in the mantle these minerals consist of more compact arrangements of the familiar building blocks of surface mineralogy: tetrahedra of SiO_4^{4-} and AlO_4^{5-} neutralized by surrounding magnesium, iron, sodium, and calcium cations. At higher pressures greater densities are achieved by abandonment of these tetrahedra, the silicates occurring in more closely packed structures in which silicon is surrounded by six oxygen anions.

The relation between these results from high-pressure mineralogy and the convective motions of plate tectonics is not well understood. The mineralogical changes have usually been formulated in terms of a rather static layered mantle, whereas plate tectonics requires an irregular and dynamic mantle. All this must be blended into the single reality of the actual earth, in which the phase changes are both causes and consequences of the thermal profiles associated with the convective flow.

The History of the Earth

It is clear that an enormous amount must be learned before we will have an adequate description of the earth's interior. Furthermore, a description of the present earth is only a portion of the burden that has been assumed by earth scientists. It is also necessary to understand the origin and history of the earth—that is, the extent to which its present physical, chemical, and dynamical state was the same or different in the past. In order to deal intelligently with many problems of the human environment, such as climatic changes, earthquakes, and changes in sea level, it is even necessary to project our understanding of the state of the earth into the future.

For several centuries geologists have painstakingly developed maps showing the distribution of the rock formations of the earth's surface. Development of a stratigraphic time scale based on extinct organisms preserved as fossils in sedimentary rocks has permitted worldwide correlation of the events of geological history extending back about 600 million years. More recently, methods of dating commonly occurring rocks by use of natural radioisotopes of uranium, thorium, rubidium, potassium, and samarium have enabled us to extend this time scale over most of the 4500 million years of earth history. Before the demonstration of plate tectonics, these dating methods permitted delineation of the trends of ancient mountain ranges, the location of volcanic regions, and the distinction between stable and mobile regions of the continents in the past. However, a usable theoretical basis was missing. For example, when a 1100-million-year mountain belt was traced from eastern Canada through the eastern United States and south to Texas, there was no basis other than naïve extrapolation for believing it ended there or continued further in a southwesterly direction. It may be hoped that our new understanding of present-day global dynamical processes will lead to solutions of problems such as this. But these solutions are not coming quickly. The most clear-cut expressions of these dynamical processes are found in the ocean basins, and the ancient ocean basins have been nearly completely consumed by these same dynamical processes. Only rarely can fragments of ancient oceanic crust be discerned, and seldom in an unambiguous way. Evidence for ancient continental rifting and continent-continent collisions are more common, but even here interpretation of the evidence is usually properly controversial. Also uncertain is the extent to which it is even legitimate to extrapolate currently observed tectonics to very early geological times. The fundamental physical and chemical laws that govern convection and mineral stability have certainly always been valid. However, until we understand exactly how and why these laws express themselves in terms of the size, number, thickness, and velocity of today's plates, we cannot be sure how similar these present plates should be to those of the past. Far enough back in time, the earth must have been different. At the time it was formed its internal production of heat by radioactivity was about three times the present value. Decay of radioactive heat sources and loss of heat by surface cooling must have resulted in considerable thermal evolution. The extent of crustal differentiation, the composition of the atmosphere, and the size of the oceans and continents have probably all evolved with time, but quantification of this evolution remains a challenge.

Moon, Planets, and the Origin of the Earth

Much attention is being given to these problems of extrapolating present observations backward in time. Considerable progress is being made toward their solution, and much more progress may be anticipated. But will it ever be possible to infer from geological, geophysical, and geochemical observations the terrestrial conditions and history all the way back to the earth's beginning?

Pessimism in answering this question is certainly justifiable. Fortunately the answer may not entirely depend on such extrapolation. Some help may be expected from the opposite direction. This is a consequence of our living in a "golden age" of planetary exploration. During the last decade study of the solar system, all the way out to Saturn, has been transformed from a branch of observational astronomy into an area of photogeology. Among other things, these planetary studies hold considerable promise of at least defining which initial boundary conditions on the history of the earth and other planets are probable and which are not.

Relatively few rocks have been preserved on the earth for more than about 2800 million years and the oldest well-dated rocks on the earth, found in West Greenland, are about 3800 million years in age. Rocks of similar, but perhaps somewhat younger age are found in isolated regions elsewhere on the earth. Why do we not find older rocks? Prior to the present epoch of lunar and planetary exploration it was not possible to give a good answer to a question of this kind. It seemed quite possible that the earth formed at a very low temperature, and that the oldest rocks found are simply relics of the first episodes of terrestrial magmatic activity and sedimentation, which may have occurred hundreds of millions of years later than the formation of the earth. Alternatively, the earth might have been completely melted at the time of its formation and remained too hot to preserve any record of times earlier than the ages of the oldest known rocks. Other alternatives can be imagined.

The rocks returned from the moon by the Apollo and Luna programs have provided the best available glimpse of the igneous record of a planetary body during the first 700 million years of solar system history. In contrast to those on the earth, volcanic rocks ranging in age from 3300 million to 3800 million years are abundant on the dark lunar maria. Even older rocks and rock fragments, commonly 3900 million to 4000 million years in age, dominate the brighter highland regions. Fragments of rocks which record a time of about 4200 million years since their last extensive heating and melting can also be found. Rare bits of older rocks are known, approaching the age of the earth and moon themselves. These very ancient rocks are interpreted as showing that the moon experienced very widespread and deep-seated igneous fractionation as early as 4400 million or 4500 million years ago. Although the details of these events remain controversial, it is important to know that even a planet as small as the moon could be very hot very early. This conclusion is underscored by the dating of basaltic meteorites at 4500 million years, these objects presumably being derived from an even smaller asteroidal parent planet. Although alternative theories for the formation of the moon and terrestrial planets remain viable, a feature common to almost all the plausible alternatives is the conclusion that if the moon was hot very early in its history, the earth must have been even hotter. Volcanism and the outgassing of volatile compounds, principally steam and carbon dioxide, must have been ubiquitous on the early earth.

Planetary exploration is also telling us other important things about the initial states of the planets and the early solar system, but we have not yet been clever enough to understand just what these things mean. For example, the first spacecraft to study Venus demonstrated that the atmosphere of our sister planet consists primarily of carbon dioxide, and the total mass of this atmosphere is very large, leading to surface atmospheric pressures about 100 times that of the earth. This massive carbon dioxide atmosphere produces a "greenhouse effect" whereby the atmosphere is relatively transparent to the short-wavelength incident solar radiation, but opaque to the longer wavelength infrared radiation which Venus radiates back into space. This effect causes its steady-state surface temperature resulting from the balance of incoming and outgoing radiation to be higher than that of the earth, about 450°C in contrast to the earth's 20°C. At the higher temperatures of Venus the earth would also have a massive atmosphere consisting mainly of water and carbon dioxide derived from carbonate rocks, and the absorption of infrared radiation would cause a similar greenhouse effect, maintaining the high temperature.

Thus it looks as though planets the size and composition of the earth and Venus can exist in one or the other of two modes. In the one mode their surface temperatures and atmospheric pressures could be low like those of the earth, and most of the volatile gases would be in the solid or liquid form. If, however, slightly different initial conditions caused the surface to be a bit hotter, enough volatile material would evaporate and the greenhouse effect would heat the surface even more, leading to further evaporation and rapid evolution into a Venus-like atmosphere. Thus one might suspect that there is no very fundamental difference between the atmosphere of the earth and Venus. The lack of water on Venus may be attributable to a greater loss of hydrogen following the breakdown of water by solar ultraviolet radiation, this being a result of the different thermal structure of the Venus atmospheric mode.

However, more recent measurements by the Venera and Pioneer probes of the Venus atmosphere reveal a more fundamental and surprising difference between the atmospheres of these two planets of similar size. With the exception of helium, which is easily lost into space, Venus has about 100 times as much of the completely volatile noble gases as does the earth. The only reason argon is a moderately abundant constituent of the earth's atmosphere is that one of the argon isotopes (^{40}Ar) is formed by the radioactive decay of a potassium isotope. In contrast, if these measurements are correct, Venus has a much more abundant supply of ordinary argon, not derived from potassium, as well as neon, and perhaps krypton and xenon as well. There is no way the greenhouse effect can cause this difference, because unlike carbon dioxide and water, these noble gases cannot be condensed at the surface temperature of the earth. It seems almost certain that this excess supply of noble gas has been present on Venus ever since the planet was formed, but current theories of planet formation fail to give a straightforward explanation of why Venus should have received these gases, and not the earth. Several explanations have been proposed, but they have at first the suspicious aspect of being ad hoc. However, if it turns out that these very special explanations are the only possible ones, these new observations will place very stringent constraints on the manner by which these planets could have formed.

The history of volatile elements on Mars is also puzzling. The total atmospheric pressure on Mars is very slight, less than 1 percent that of the earth. Like the atmosphere of Venus, it consists mainly of carbon dioxide, with only traces of water vapor. Yet photographs of the martian surface by Mariner and Viking spacecraft show that its surface seems to have been eroded by running water, probably associated with sudden catastrophic flooding. The present atmospheric pressure and temperature on Mars is too low to permit the existence of liquid water on the surface, and martian water must be almost entirely confined to its permanent polar caps and to ground ice in its permafrost soil. Could the temperature on Mars have been warm enough to produce a higher pressure water-rich atmosphere with running water on its surface? Celestial mechanical studies of the way the orbit of Mars and the orientation of its rotation axis in its orbit varies with time suggest that the surface temperature of Mars may vary considerably on a time scale of millions of years. Nevertheless it is difficult to understand quantitatively how this effect could be large enough to allow liquid water to produce the observed erosional phenomena. Perhaps the erosional features were caused by lava with a viscosity so low that it could mimic the effect of running water. Or maybe some qualitatively new insight is required to solve this mystery.

The atmosphere of Mars contains very little argon produced by the radioactive decay of potassium. Although it is possible that Mars contains much less potassium than the earth, there is no good reason to believe this to be true. In fact conventional wisdom would hold that being further from the sun, Mars should contain a higher proportion of the relatively volatile element potassium, which is depleted on the earth by about a factor of 10, relative to its average solar system abundance. A more likely explanation is that most of this argon is still locked up in the interior of Mars and that its heat engine runs more slowly than that of the earth, probably because of a lower initial temperature.

Another major finding of this epochal period of space exploration is that the most ancient lunar rocks are never found as large fresh pieces of frozen lava like the basalts of the dark mare regions. They are invariably found as naturally welded fragments in complex breccias. Both their isotopic chemistry and appearance indicate they are the end products of the intense bombardment of extralunar bodies that produced the densely cratered terrain of the lunar highlands. Radioactive dating shows that this heavy bombardment ceased quite suddenly about 3900 million years ago. The mare surfaces that postdate this age are relatively lightly cratered, and the rocks themselves unmetamorphosed by the shock pressures that accompany hypervelocity impact.

There is no way the earth could have avoided a similar cratering history. Thus lunar studies show that the absence on earth of primordial rocks is twofold in origin: very extensive magmatic and geological activity of internal origin, combined with massive external impact metamorphism of all primordial surfaces. Dynamical studies show that if the impacting bodies were in heliocentric, rather than geocentric orbits, all of the terrestrial planets experienced a similar bombardment history.

Indeed, the most obvious characteristic of the photographs of Mercury and Mars obtained from spacecraft is the cratering record produced by this ancient bombardment. The abrupt decline in impact rate of 3900 million years may constitute a "marker horizon" throughout the inner solar system, separating earlier events from later ones.

These same photographs, and radar studies of Venus, show that all these planets have volcanoes, and thus internal heat engines as well. The magnitude of internal volcanic activity since that time varies from one planet to another. Much effort will be required to understand how these other heat engines operate, and to what extent their differences are a consequence of planetary mass, chemical composition, distance from the sun, and initial boundary conditions. Our plate tectonic world is obviously not the only possible world. Neither the moon, Mars, nor Mercury show evidence for such plate motion.

Preliminary radar data obtained from Pioneer Venus suggest that plate tectonics is also absent on that planet. Entirely different tectonic styles may be present on the large Galilean satellites of Jupiter, photographed by the two Voyager flybys. Most dramatic is the violent, sulfur-rich volcanism on the innermost large satellite Io, driven by tidal energy pumped into the satellite by its giant neighbor planet.

Major new insights into conditions early in the history of the solar system are also being provided by research on meteorites (Fig. 5). Of particular importance are the unequilibrated objects, such as the Allende meteorite, and the carbonaceous chondrites, which preserve the record of heterogeneity of the early solar system, and may even provide a glimpse into our presolar history.

Until recently, it was fashionable for practical geologists to regard discussions of the formation of the earth and solar system as of little relevance to their work. This lack of relevance permitted theoretically minded scientists to speculate very freely, perhaps even irresponsibly, about these things. These times are over. The unity and interdependence of earth and planetary science, as well as of stellar and even galactic astronomy, are now apparent. Perhaps the greatest challenge during the next decade will be to integrate these sciences, while at the same time preserving the rigorous internal professional discipline that is the hard-won result of many generations of tradition.

Additional Readings

1. Geophysics Study Committee, *Continental Tectonics* (Geophysics Research Board, National Research Council, National Academy of Sciences, Washington, D.C., 1980).
2. U.S. National Committee on Geochemistry, *Orientations in Geochemistry* (National Research Council, National Academy of Sciences, Washington, D.C., 1973).
3. U.S. Geodynamics Committee, *Geodynamics in the 1980's* (Geophysics Research Board, National Research Council, National Academy of Sciences, Washington, D.C., 1980).
4. R. K. Bambach, C. R. Scotese, A. M. Ziegler, "Before Pangea: The geographies of the paleozoic world," *Am. Sci.* **68**, 26 (January-February, 1980).
5. J. G. Sclater and C. Tapscott, "The history of the Atlantic," *Sci. Am.* **240**, 156 (June 1979).
6. T. H. Jordan, "The deep structure of continents," *ibid.* **240**, 92 (January 1979).
7. B. Marsh, "Island arc volcanism," *Am. Sci.* **67**, No. 2 (March-April, 1979).
8. D. L. Peck, T. L. Wright, R. W. Decker, "The lava lakes of Kilauea," *Sci. Am.* **241**, 114 (October 1979).
9. C. G. Chase, E. M. Herron, W. R. Normak, "Plate tectonics: Commotion in the ocean and continental consequences," *Annu. Rev. Earth Planet Sci.* **3**, 271 (1975).
10. K. Burke, "The Aulacogens and continental breakup," *ibid.* **5**, 371 (1977).
11. E. Bullard, "The emergence of plate tectonics: A personal view," *ibid.* **3**, 1 (1975).
12. *The Future of Scientific Ocean Drilling*, Report by an Ad Hoc Subcommittee of the JOIDES (Joint Oceanographic Institutions for Deep Earth Sampling) Executive Committee, Seattle, Washington, 1977).
13. U.S. Geodynamics Committee, *Continental Scientific Drilling Program* (Geophysics Research Board, National Research Council, National Academy of Sciences, Washington, D.C., 1979).
14. Panel on Continental Margins, *Continental Margins* (Ocean Sciences Board, National Research Council, National Academy of Sciences, Washington, D.C., 1979).
15. U.S. National Report to IUGG, *Rev. Geophys. Space Phys.* **17**, Nos. 2-7 (1979).
16. U.S. Geodynamics Committee, *U.S. Program for the Geodynamics Project: Scope and Objectives* (Geophysics Research Board, National Research Council, National Academy of Sciences, Washington, D.C., 1973).
17. W. K. Hartmann, *Astronomy: The Cosmic Journey* (Wadsworth, Belmont, Calif., 1978).

Mathematics

Saunders Mac Lane

Today, active mathematical research goes on in a bewildering variety of interlocking fields and subfields (*1*). Some of these fields are long-established and relatively concrete ones, like geometry, number theory, or differential equations.

Summary. Current research in mathematics involves a wide variety of interlocking ideas, old and new. For example, results about the curves and surfaces defined by polynomial equations, as in algebraic geometry, appear in the study of solitary waves and also in the gauge theories in physics. Centuries-old problems in number theory have been solved, while others have been revealed as insoluble. The classification of all finite simple groups is nearly achieved (and the full treatment will be voluminous); the representation of groups aids in their application to the study of symmetry. These developments, and many others, attest to the vitality of mathematics.

Others are new and apparently abstract, such as functional analysis or algebraic K-theory. In almost all of these topics, one has today a variety of exciting developments, some closely connected with applications, as in the case of the theory of linear systems. Other active fields, such as algebraic geometry, seem far removed from applications. However, startling connections between topics do appear, as in the current discovery that notions from algebraic geometry can be used in solving nonlinear partial differential equations (the solitons) and in the gauge theories, which are now of central interest in theoretical physics (in relation to instantons and the Yang-Mills equations). Similar surprising interactions occur between apparently very different branches of pure mathematics—as illustrated by the discussion of the Smith conjecture below, or by the examples in *Mathematics Today* (*2*).

Recent years have seen extraordinary progress in mathematics, as measured either by the large number of famous problems and difficult problems which have been solved (*2, 3*) or by the remarkable number of new concepts which have been developed. I will first give some examples of problems solved and of concepts developed. Then I will try to give more detailed examples of some of the new fields of mathematical research. First I consider some of the old problems recently solved.

Mathematical problems arise naturally, but they can be very recalcitrant, and sometimes literally insoluble. Thus Fourier series have been used to describe periodic phenomena for nearly 200 years. The basic convergence theorem for these series was conjectured about 1910 by the Russian N. Luzin, but was established just in 1966 by Carleson (*4*) (convergence "almost everywhere," given a function in L^2).

In the early 1800's Gauss showed that the complex integers $m + ni$, with $i = \sqrt{-1}$ and m and n whole numbers, could be decomposed uniquely into primes. Gauss also found eight other such cases of unique decomposition for integers $m + n\sqrt{-d}$, where d is positive. It was conjectured that there might be just one more case (making ten in all), but Heegner (*5*), Baker (*6*), and Stark (*7*) proved that there are only nine such cases. A recurring problem is that of determining whether or not such an explicit number is irrational (like $\sqrt{2}$ or $\sqrt[3]{7}$) or transcendental (like e or π). There has been much recent progress here; for example, for the famous Riemann zeta function ζ. R. Apéry succeeded in 1978 in proving that $\zeta(3)$ is irrational (*8*).

The four-color conjecture first appeared 100 years ago; it asserted that on any map of different countries on the globe, it is possible to color the land of each country, using at most four different colors, in such a way that no two countries with a common boundary have the same color. Numerous attempts to prove this failed, but now Appel and Haken (*9*) have established this theorem by using a classical method of attack supplemented by a massive calculation, of about 2000 special configurations, on a computer.

Geometry has made great strides in the classification of closed manifolds. In two dimensions, these manifolds are the closed surfaces such as the sphere, the torus (inner tube), the surface of a pretzel with two holes, three holes, and so on. In fact, this list can be proved to be a complete classification for two-sided sur-

The author is Max Mason Distinguished Service Professor of Mathematics at the University of Chicago, Chicago, Illinois 60637.