SPECIAL EDITION

SCIENTIFIC AMERICAN

WWW.SCIAM.COM

Our Ever Changing EARTH

Moving **Mountains**

Power of the **Geodynamo**

Nomadic **Continents**

Surface **Sculpting** from Within

Secret **Earthquakes**

Revealed: **Seafloor Vistas**

letter from the editor

A Deeper Understanding


KENN BROWN

IT'S SO BORING, the usual human's-eye view. Seasons come and go, but terra firma itself never varies. Even an earthquake or a mudslide seems like a random incident unconnected to any larger or more complex patterns.

But put on the lenses of a geologist and take another look. Reading the stories imprinted on the rocks and crystals gives scientists the ability to examine our world as it has evolved over millions, even billions, of years. From this vantage point, it is easy to see that Earth has been—and continues to be—a lively cauldron of change. Just as stop-action photography shows how buds burst into flower, geology gives us a picture of a living, changing planet. It even has a heartbeat of sorts, in pulsed releases of vast amounts of inner heat. The most recent occurred when dinosaurs still roamed, as Roger L. Larson discusses in "The Mid-Cretaceous Superplume Episode," on page 22.

As scientists draw on records of past tremors, it becomes clear that earthquakes, which seemed to be isolated, can interact with one another. In "Earthquake Conversations," on page 82, Ross S. Stein explains how these surprising interconnected effects could give nations, cities and individuals new power to evaluate local vulnerability to temblors.

In another place where taking the long view alters our everyday notions, we learn that erosion—long familiar as the great leveler—has also had a hand in shaping the tallest peaks, the Himalayas. Eroding land, by reducing weight, accelerates tectonic processes, creating an uplift. Turn to page 74 for "How Erosion Builds Mountains," by Nicholas Pinter and Mark T. Brandon.

Deep under the morphing surface, the interior simmers. The churning mantle creates and powers the primary magnetic field that surrounds the sphere. The polarity of this field reverses every so often, which has long puzzled researchers. "Probing the Geodynamo," by Gary A. Glatzmaier and Peter Olson, on page 28, tells of new, intriguing clues about how the next reversal may begin.

So join us for a jaunt in geology's rock-encrusted time machine. The articles in this special edition promise a rare look inside the mysterious and little-appreciated underfoot activities of the world we all call home.

Mariette DiChristina
Executive Editor
Scientific American
editors@sciam.com

Contents 2005

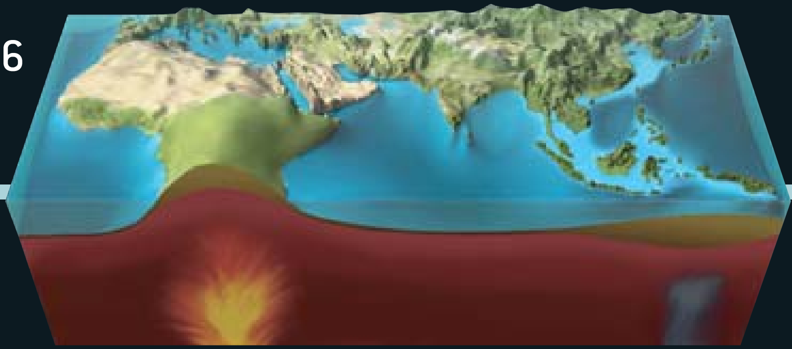
SCIENTIFIC AMERICAN Volume 15, Number 2

Our Ever Changing Earth

4

28

56

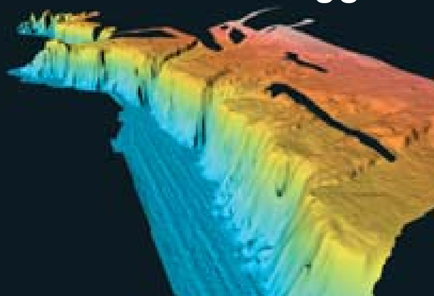


FORCES THAT SHAPE AND SHAKE EARTH

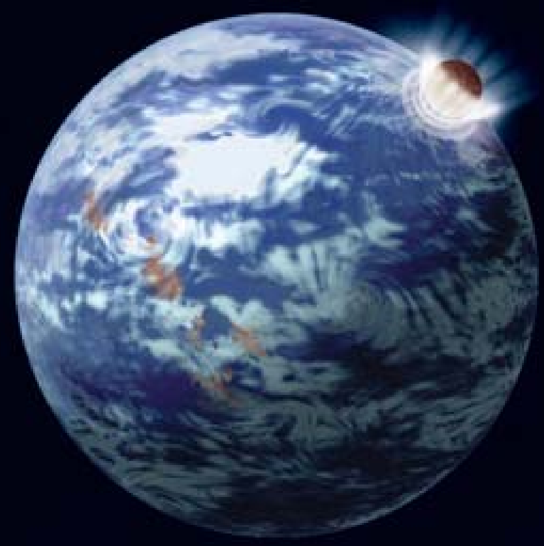
74

Cover illustration by Kenn Brown

50



Scientific American Special (ISSN 1048-0943), Volume 15, Number 2, 2005, published by Scientific American, Inc., 415 Madison Avenue, New York, NY 10017-1111. Copyright © 2005 by Scientific American, Inc. All rights reserved. No part of this issue may be reproduced by any mechanical, photographic or electronic process, or in the form of a phonographic recording, nor may it be stored in a retrieval system, transmitted or otherwise copied for public or private use without written permission of the publisher. Canadian BN No. 127387652RT; QST No. Q1015332537. To purchase additional quantities: U.S., $10.95 each; elsewhere, $13.95 each. Send payment to Scientific American, Dept. ES, 415 Madison Avenue, New York, NY 10017-1111. Inquiries: fax 212-355-0408 or telephone 212-451-8890. Printed in U.S.A.



The Evolution of Earth

By Claude J. Allègre and Stephen H. Schneider

The evolution of this planet and its atmosphere gave rise to life, which shaped Earth's subsequent development. Our future lies in interpreting this geologic past and considering what changes—good and bad—may lie ahead

EARTH SEEN FROM SPACE has changed dramatically. One hundred million years after it had formed—some four billion years ago—the planet was probably undergoing meteor bombardment (*left*). At this time, it may have been studded with volcanic islands and shrouded by an atmosphere laden with carbon dioxide and heavy with clouds. Three billion years ago its face may have been obscured by an orange haze of methane, the product of early organisms (*center*). Today clouds, oceans and continents are clearly discernible (*right*). This illustration was prepared with the help of James F. Kasting of Pennsylvania State University.

Like the lapis lazuli gem it resembles, the blue, cloud-enveloped planet that we recognize immediately from satellite pictures seems remarkably stable.

Continents and oceans, encircled by an oxygen-rich atmosphere, support familiar life-forms. Yet this constancy is an illusion produced by the human experience of time. Earth and its atmosphere are continuously altered. Plate tectonics shift the continents, raise mountains and move the ocean floor while processes not fully understood alter the climate.

Such constant change has characterized Earth since its beginning some 4.5 billion years ago. From the outset, heat and gravity shaped the evolution of the planet. These forces were gradually joined by the global effects of the emergence of life. Exploring this past offers us the only possibility of understanding the origin of life and, perhaps, its future.

Scientists used to believe the rocky planets, including Earth, Mercury, Venus and Mars, were created by the rapid gravitational collapse of a dust cloud, a deflation giving rise to a dense orb. In the 1960s the Apollo space program changed this view. Studies of moon craters revealed that these gouges were caused by the impact of objects that were in great abundance about 4.5 billion years ago. Thereafter, the number of impacts appeared to have quickly decreased. This observation rejuvenated the theory of accretion postulated by Otto Schmidt. The Russian geophysicist had suggested in 1944 that planets grew in size gradually, step by step.

According to Schmidt, cosmic dust lumped together to form particulates, particulates became gravel, gravel became small balls, then big balls, then tiny planets, or planetesimals, and, finally, dust became the size of the moon. As the planetesimals became larger, their numbers decreased. Consequently, the number of collisions between planetesimals, or meteorites, decreased. Fewer items available for accretion meant that it took a long time to build up a large planet. A calculation made by George W. Wetherill of the Carnegie Institution of Washington suggests that about 100 million years could pass between the formation of an object measuring 10 kilometers in diameter and an object the size of Earth.

The process of accretion had significant thermal consequences for Earth, consequences that forcefully directed its evolution. Large bodies slamming into the planet produced immense heat in its interior, melting the cosmic dust found there. The resulting furnace—situated some 200 to 400 kilometers underground and called a magma ocean—was active for millions of years, giving rise to volcanic eruptions. When Earth was young, heat at the surface caused by volcanism and lava flows from the interior was intensified by the constant bombardment of huge objects, some of them perhaps the size of the moon or even Mars. No life was possible during this period.

Beyond clarifying that Earth had formed through accretion, the Apollo program compelled scientists to try to reconstruct the subsequent temporal and physical development of the early Earth. This undertaking had been considered impossible by founders of geology, including Charles Lyell, to whom the following phrase is attributed: No vestige of a beginning, no prospect for an end. This statement conveys the idea that the young Earth could not be recreated, because its remnants were destroyed by its very activity. But the development of isotope geology in the 1960s had rendered this view obsolete. Their imaginations fired by Apollo and the moon findings, geochemists began to apply this technique to understand the evolution of Earth.

Dating rocks using so-called radioactive clocks allows geologists to work on old terrains that do not contain fossils. The hands of a radioactive clock are isotopes—atoms of the same element that have different atomic weights—and geologic time is measured by the rate of decay of one isotope into another [see "The Earliest History of the Earth," by Derek York; SCIENTIFIC AMERICAN, January 1993]. Among the many clocks, those based on the decay of uranium 238 into lead 206 and of uranium 235 into lead 207 are special. Geochronologists can determine the age of samples by analyzing only the daughter product—in this case, lead—of the radioactive parent, uranium.

PANNING FOR ZIRCONS

ISOTOPE GEOLOGY has permitted geologists to determine that the accretion of Earth culminated in the differentiation of the planet: the creation of the core—the source of Earth's magnetic field—and the beginning of the atmosphere. In 1953 the classic work of Claire C. Patterson of the California Institute of Technology used the uranium-lead clock to establish an age of 4.55 billion years for Earth and many of the meteorites that formed it. In the early 1990s, however, work by one of us (Allègre) on lead isotopes led to a somewhat new interpretation.

As Patterson argued, some meteorites were indeed formed about 4.56 billion years ago, and their debris constituted Earth. But Earth continued to grow through the bombardment of planetesimals until some 120 million to 150 million years later. At that time—4.44 billion to 4.41 billion years ago—Earth began to retain its atmosphere and create its core. This possibility had already been suggested by Bruce R. Doe and Robert E. Zartman of the U.S.

JACK HARRIS *Visual Logic* AND DIMITRY SCHIDLOVSKY (*preceding pages*)

Geological Survey in Denver two decades ago and is in agreement with Wetherill's estimates.

The emergence of the continents came somewhat later. According to the theory of plate tectonics, these landmasses are the only part of Earth's crust that is not recycled and, consequently, destroyed during the geothermal cycle driven by the convection in the mantle. Continents thus provide a form of memory because the record of early life can be read in their rocks. Geologic activity, however, including plate tectonics, erosion and metamorphism, has destroyed almost all the ancient rocks. Very few fragments have survived this geologic machine.

Nevertheless, in recent decades, several important finds have been made, again using isotope geochemistry. One group, led by Stephen Moorbath of the University of Oxford, discovered terrain in West Greenland that is between 3.7 billion and 3.8 billion years old. In addition, Samuel A. Bowring of the Massachusetts Institute of Technology explored a small area in North America—the Acasta gneiss—that is thought to be 3.96 billion years old.

Ultimately, a quest for the mineral zircon led other researchers to even more ancient terrain. Typically found in continental rocks, zircon is not dissolved during the process of erosion but is deposited in particle form in sediment. A few pieces of zircon can therefore survive for billions of years and can serve as a witness to Earth's more ancient crust. The search for old zircons started in Paris with the work of Annie Vitrac and Joël R. Lancelot, later at the University of Marseille and now at the University of Nîmes, respectively, as well as with the efforts of Moorbath and Allègre. It was a group at the Australian National University in Canberra, directed by William Compston, that was finally successful. The team discovered zircons in western Australia that were between 4.1 billion and 4.3 billion years old.

Zircons have been crucial not only for understanding the age of the continents but for determining when life first appeared. The earliest fossils of undisputed age were found in Australia and South Africa. These relics of blue-green algae are about 3.5 billion years old. Manfred Schidlowski of the Max Planck Institute for Chemistry in Mainz studied the Isua formation in West Greenland and argued that organic matter existed as long ago as 3.8 billion years. Because most of the record of early life has been destroyed by geologic activity, we cannot say exactly when it first appeared—perhaps it arose very quickly, maybe even 4.2 billion years ago.

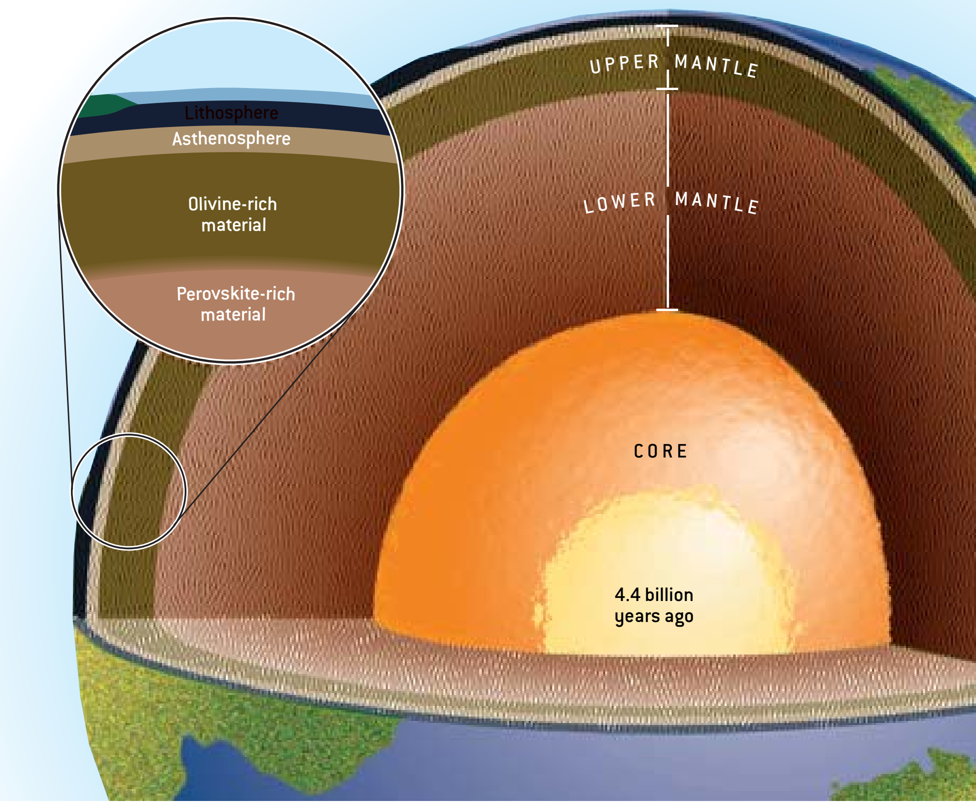
STORIES FROM GASES

ONE OF THE MOST important aspects of the planet's evolution is the formation of the atmosphere, because it is this assemblage of gases that allowed life to crawl out of the oceans and to be sustained. Researchers have hypothesized since the 1950s that the terrestrial atmosphere was created by gases emerging from the interior of the planet. When a volcano spews gases, it is an example of the continuous outgassing, as it is called, of Earth. But scientists have questioned whether this process occurred suddenly—about 4.4 billion years ago when the core differentiated—or whether it took place gradually over time.

To answer this question, Allègre and his colleagues studied the isotopes of rare gases. These gases—including helium, argon and xenon—have the peculiarity of being chemically inert, that is, they do not react in nature with other elements. Two of them are particularly important for atmospheric studies: argon and xenon. Argon has three

HOW EARTH GOT ITS CORE

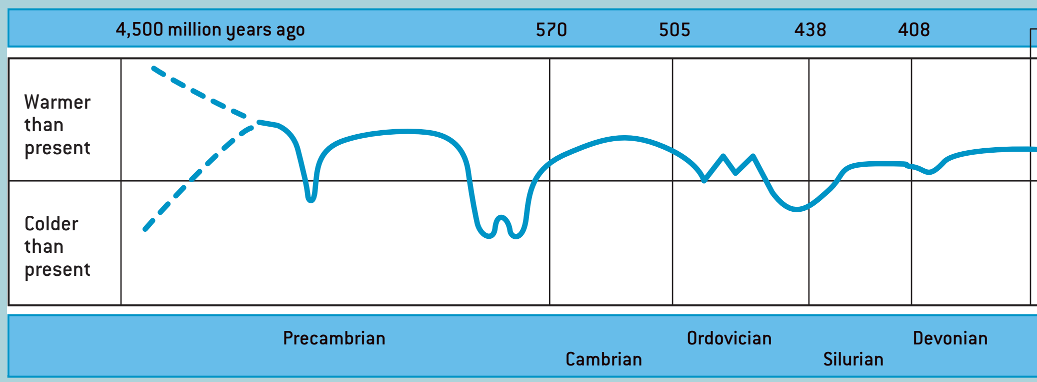
THE DIFFERENTIATION of the planet took place quite quickly after Earth was formed by the accretion of cosmic dust and meteorites. About 4.4 billion years ago the core—which, with the mantle, drives the geothermal cycle, including volcanism—appeared; gases emerging from the interior of the planet also gave rise to a nascent atmosphere. Somewhat later, although the issue has not been entirely resolved, it seems that continental crust formed as the various elements segregated into different depths.



IAN WORPOLE

LIFE EVOLVES AS CLIMATE CYCLES

CLIMATE FLUCTUATIONS are apparent over time. Although Earth's early temperature record is quite uncertain, good estimates can be made starting 400 million years ago, when fossils were more abundantly preserved. As climate shifted, so did life—suggesting feedback between the two. The dates of these evolutions remain unclear as well, but their order is more apparent. First a primordial soup formed, then primitive organisms, such as algae, stromatolites and jellyfish, arose; spiny fish were followed by the ichthyostega, perhaps the first creature to crawl from ocean onto land. The rest of the story is well known: dinosaurs appeared and died out, their place taken by mammals.



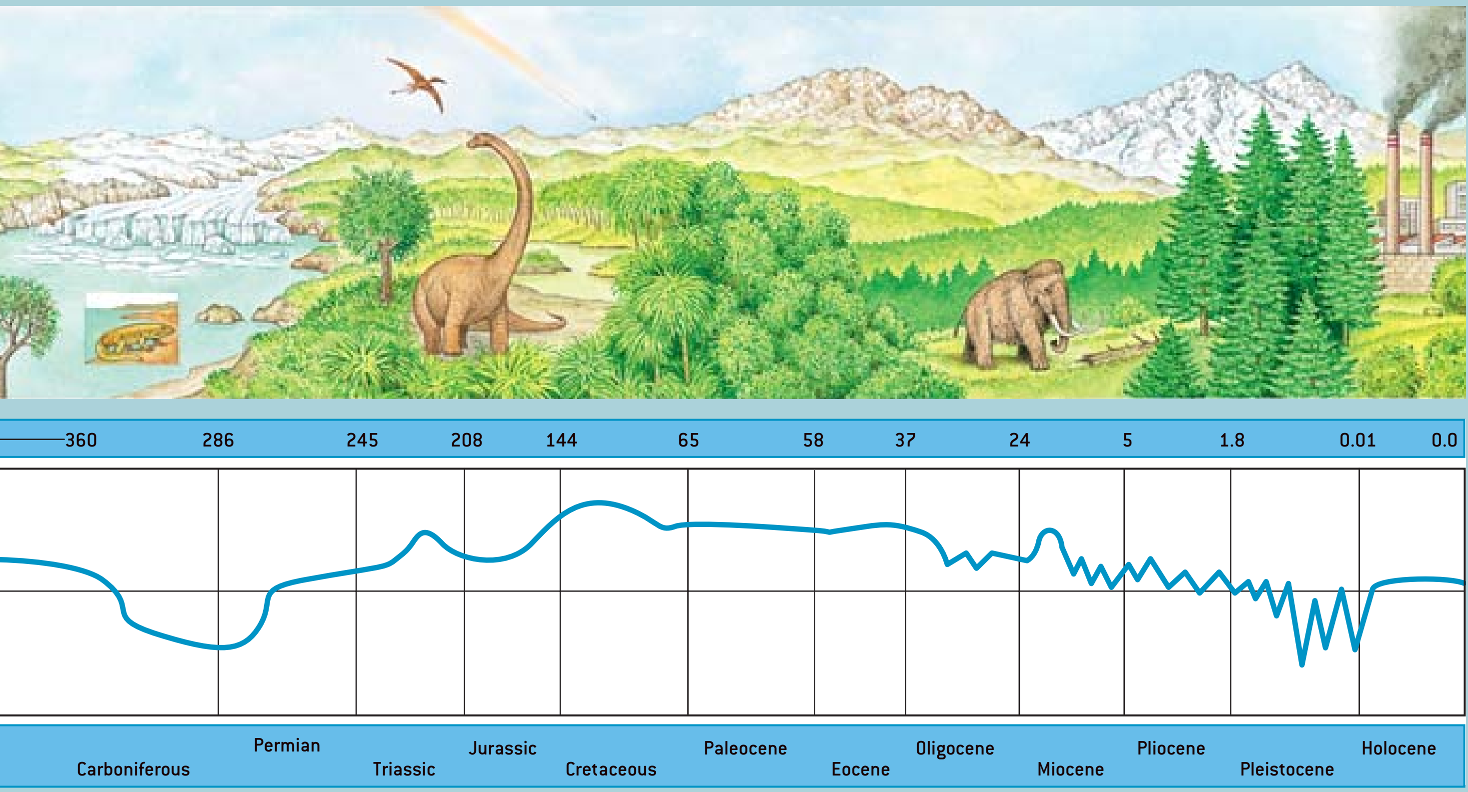
ROBERTO OSTI (*illustration*); LAURIE GRACE (*graph*)

isotopes, of which argon 40 is created by the decay of potassium 40. Xenon has nine, of which xenon 129 has two different origins. Xenon 129 arose as the result of nucleosynthesis before Earth and solar system were formed. It was also created from the decay of radioactive iodine 129, which does not exist on Earth anymore. This form of iodine was present very early on but has died out since, and xenon 129 has grown at its expense.

Like most couples, both argon 40 and potassium 40 and xenon 129 and iodine 129 have stories to tell. They are excellent chronometers. Although the atmosphere was formed by the outgassing of the mantle, it does not contain any potassium 40 or iodine 129. All argon 40 and xenon 129, formed in Earth and released, are found in the atmosphere today. Xenon was expelled from the mantle and retained in the atmosphere; therefore, the atmosphere-mantle ratio of this element allows us to evaluate the age of differentiation. Argon and xenon trapped in the mantle evolved by the radioactive decay of potassium 40 and iodine 129. Thus, if the total outgassing of the mantle occurred at the beginning of Earth's formation, the atmosphere would not contain any argon 40 but would contain xenon 129.

The major challenge facing an investigator who wants to measure such ratios of decay is to obtain high concentrations of rare gases in mantle rocks because they are extremely limited. Fortunately, a natural phenomenon occurs at mid-ocean ridges during which volcanic lava transfers some silicates from the mantle to the surface. The small amounts of gases trapped in mantle minerals rise with the melt to the surface and are concentrated in small vesicles in the outer glassy margin of lava flows. This process serves to concentrate the amounts of mantle gases by a factor of 10^4 or 10^5. Collecting these rocks by dredging the seafloor and then crushing them under vacuum in a sensitive mass spectrometer allows geochemists to determine the ratios of the isotopes in the mantle. The results are quite surprising. Calculations of the ratios indicate that between 80 and 85 percent of the atmosphere was outgassed during Earth's first one million years; the rest was released slowly but constantly during the next 4.4 billion years.

The composition of this primitive atmosphere was most certainly dominated by carbon dioxide, with nitrogen as the second most abundant gas. Trace amounts of methane, ammonia, sulfur dioxide and hydrochloric acid were also present, but there was no oxygen. Except for the presence of abundant water, the atmosphere was similar to that of Venus or Mars. The details of the evolution of the original atmosphere are debated, particularly because we do not know how strong the sun was at that time. Some facts, however, are not disputed. It is evident that carbon dioxide played a crucial role. In addition, many scientists believe the evolving atmosphere contained sufficient quantities of gases such as ammonia and methane to give rise to organic matter.



Still, the problem of the sun remains unresolved. One hypothesis holds that during the Archean eon, which lasted from about 4.5 billion to 2.5 billion years ago, the sun's power was only 75 percent of what it is today. This possibility raises a dilemma: How could life have survived in the relatively cold climate that should accompany a weaker sun? A solution to the faint early sun paradox, as it is called, was offered by Carl Sagan and George Mullen of Cornell University in 1970. The two scientists suggested that methane and ammonia, which are very effective at trapping infrared radiation, were quite abundant. These gases could have created a super-greenhouse effect. The idea was criticized on the basis that such gases were highly reactive and have short lifetimes in the atmosphere.

WHAT CONTROLLED CO_2?

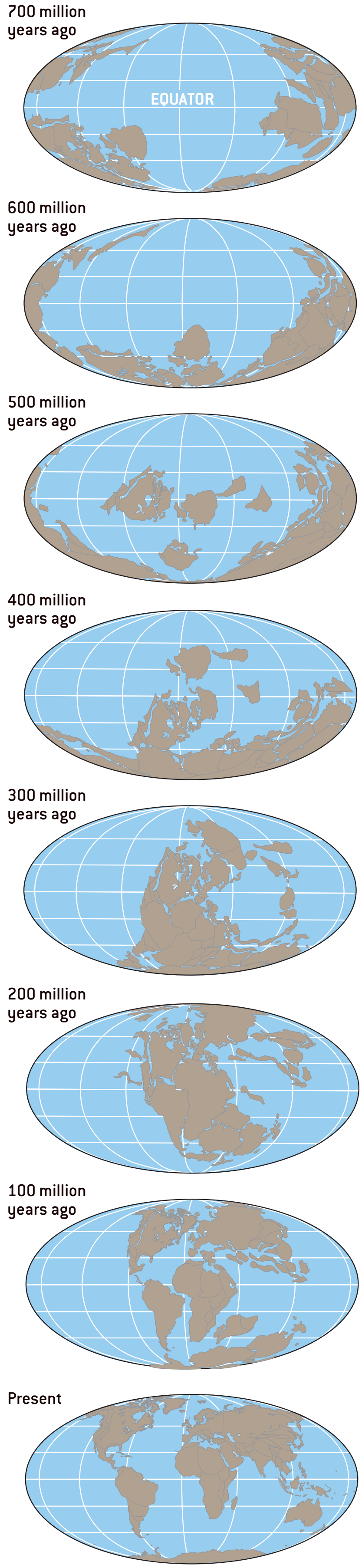
IN THE LATE 1970S Veerabhadran Ramanathan, now at the Scripps Institution of Oceanography, and Robert D. Cess and Tobias Owen of Stony Brook University proposed another solution. They postulated that there was no need for methane in the early atmosphere because carbon dioxide was abundant enough to bring about the super-greenhouse effect. Again this argument raised a different question: How much carbon dioxide was there in the early atmosphere? Terrestrial carbon dioxide is now buried in carbonate rocks, such as limestone, although it is not clear when it became trapped there. Today calcium carbonate is created primarily during biological activity; in the Archean eon, carbon may have been primarily removed during inorganic reactions.

The rapid outgassing of the planet liberated voluminous quantities of water from the mantle, creating the oceans and the hydrologic cycle. The acids that were probably present in the atmosphere eroded rocks, forming carbonate-rich rocks. The relative importance of such a mechanism is, however, debated. Heinrich D. Holland of Harvard University believes the amount of carbon dioxide in the atmosphere rapidly decreased during the Archean and stayed at a low level.

Understanding the carbon dioxide content of the early atmosphere is pivotal to understanding climatic control. Two conflicting camps have put forth ideas on how this process works. The first group holds that global temperatures and carbon dioxide were controlled by inorganic geochemical feedbacks; the second asserts that they

THE AUTHORS

CLAUDE J. ALLÈGRE and *STEPHEN H. SCHNEIDER* study various aspects of Earth's geologic history and its climate. Allègre is professor at the University of Paris and directs the department of geochemistry at the Paris Geophysical Institute. He is a foreign member of the National Academy of Sciences. Schneider is professor in the department of biological sciences at Stanford University and co-director of the Center for Environmental Science and Policy. He was honored with a MacArthur Prize Fellowship in 1992 and was elected to membership in the National Academy of Sciences in 2002.



CONTINENTAL SHIFT has altered the face of the planet for nearly a billion years, as can be seen in the differences between the positions of the continents that we know today and those of 700 million years ago. Pangaea, the superaggregate of early continents, came together about 200 million years ago and then promptly, in geologic terms, broke apart.

IAN WORPOLE; SOURCE: CHRISTOPHER R. SCOTESE *University of Texas at Arlington*

were controlled by biological removal.

James C. G. Walker, James F. Kasting and Paul B. Hays, then at the University of Michigan at Ann Arbor, proposed the inorganic model in 1981. They postulated that levels of the gas were high at the outset of the Archean and did not fall precipitously. The trio suggested that as the climate warmed, more water evaporated, and the hydrologic cycle became more vigorous, increasing precipitation and runoff. The carbon dioxide in the atmosphere mixed with rainwater to create carbonic acid runoff, exposing minerals at the surface to weathering. Silicate minerals combined with carbon that had been in the atmosphere, sequestering it in sedimentary rocks. Less carbon dioxide in the atmosphere meant, in turn, less of a greenhouse effect. The inorganic negative feedback process offset the increase in solar energy.

This solution contrasts with a second paradigm: biological removal. One theory advanced by James E. Lovelock, an originator of the Gaia hypothesis, assumed that photosynthesizing microorganisms, such as phytoplankton, would be very productive in a high carbon dioxide environment. These creatures slowly removed carbon dioxide from the air and oceans, converting it into calcium carbonate sediments. Critics retorted that phytoplankton had not even evolved for most of the time that Earth has had life. (The Gaia hypothesis holds that life on Earth has the capacity to regulate temperature and the composition of Earth's surface and to keep it comfortable for living organisms.)

In the early 1990s Tyler Volk of New York University and David W. Schwartzman of Howard University proposed another Gaian solution. They noted that bacteria increase carbon dioxide content in soils by breaking down organic matter and by generating humic acids. Both activities accelerate weathering, removing carbon dioxide from the atmosphere. On this point, however, the controversy becomes acute. Some geochemists, including Kasting, now at Pennsylvania State University, and Holland, postulate that while life may account for some carbon dioxide removal after the Archean, inorganic geochemical processes can explain most of the sequestering. These researchers view life as a rather weak climatic stabilizing mechanism for the bulk of geologic time.

OXYGEN FROM ALGAE

THE ISSUE OF CARBON remains critical to how life influenced the atmosphere. Carbon burial is a key to the vital process of building up atmospheric oxygen concentrations—a prerequisite for the development of certain lifeforms. In addition, global warming is taking place now as a result of humans releasing this carbon. For one billion or two billion years, algae in the oceans produced oxygen. But because this gas is highly reactive and because there were many reduced minerals in the ancient oceans—iron, for example, is easily oxidized—much of the oxygen produced by living creatures simply got used up before it could reach the atmosphere, where it would have encountered gases that would react with it.

Even if evolutionary processes had given rise to more complicated lifeforms during this anaerobic era, they would have had no oxygen. Furthermore, unfiltered ultraviolet sunlight would have likely killed them if they left the ocean. Researchers such as Walker and Preston Cloud, then at the University of California at Santa Barbara, have suggested that only about two billion years ago, after most of the reduced minerals in the sea were oxidized, did atmospheric oxygen accumulate. Between one billion and two billion years ago oxygen reached current levels, creating a niche for evolving life.

By examining the stability of certain minerals, such as iron oxide or uranium oxide, Holland has shown that the oxygen content of the Archean atmosphere was low before two billion years ago. It is largely agreed that the present-day oxygen content of 20 percent is the result of photosynthetic activity. Still, the question is whether the oxygen content in the atmosphere increased gradually over time or suddenly. Recent studies indicate that the increase of oxygen started abruptly between 2.1 billion and 2.03 billion years ago and that the present situation was reached 1.5 billion years ago.

The presence of oxygen in the atmosphere had another major benefit for an organism trying to live at or above the surface: it filtered ultraviolet radiation. Ultraviolet radiation breaks down many molecules—from DNA and oxygen to the chlorofluorocarbons that are implicated in stratospheric ozone depletion. Such energy splits oxygen into the highly unstable atomic form O, which can combine back into O_2 and into the very special molecule O_3, or ozone. Ozone, in turn, absorbs ultraviolet radiation. It was not until oxygen was abundant enough in the atmosphere to allow the formation of ozone that life even had a chance to get a roothold or a foothold on land. It is not a coincidence that the rapid evolution of life from prokaryotes (single-celled organisms with no nucleus) to eukaryotes (single-celled organisms with a nucleus) to metazoa (multicelled organisms) took place in the billion-year-long era of oxygen and ozone.

Although the atmosphere was reaching a fairly stable level of oxygen during this period, the climate was hardly uniform. There were long stages of relative warmth or coolness during the transition to modern geologic time. The composition of fossil plankton shells that lived near the ocean floor provides a measure of bottom water temperatures. The record suggests that over the past 100 million years bottom waters cooled by nearly 15 degrees Celsius. Sea levels dropped by hundreds of meters, and continents drifted apart. Inland seas mostly disappeared, and the climate cooled an average of 10 to 15 degrees C. Roughly 20 million years ago permanent ice appears to have built up on Antarctica.

About two million to three million years ago the paleoclimatic record starts to show significant expansions and contractions of warm and cold periods in 40,000-year or so cycles. This periodicity is interesting because it corresponds to the time it takes Earth to complete an oscillation of the tilt of its axis of rotation. It has long been speculated, and recently calculated, that known changes in orbital geometry could alter the amount of sunlight coming in between winter and summer by about 10 percent or so and could be responsible for initiating or ending ice ages.

THE WARM HAND OF MAN

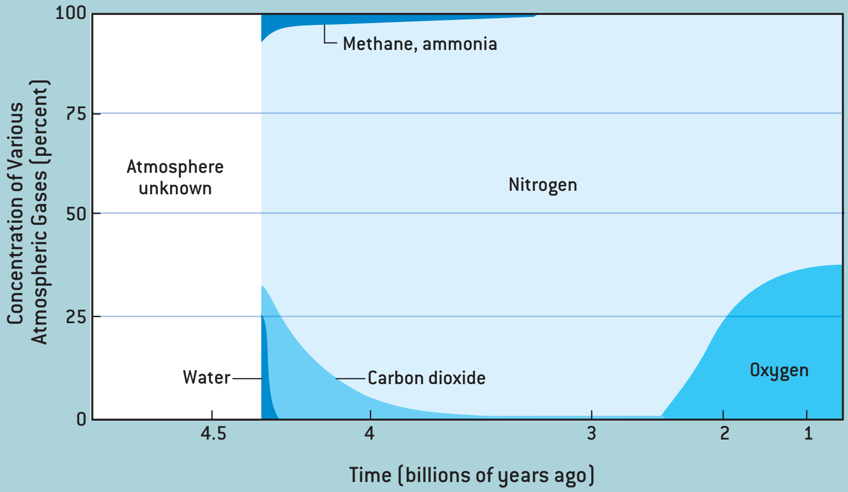
MOST INTERESTING and perplexing is the discovery that between 600,000 and 800,000 years ago the dominant cycle switched from 40,000-year periods to 100,000-year intervals with very large fluctuations. The last major phase of glaciation ended about 10,000 years ago. At its height 20,000 years ago, ice sheets about two kilometers thick covered much of northern Europe and North America. Glaciers expanded in high plateaus and mountains throughout the world. Enough ice was locked up on land to cause sea levels to drop more than 100 meters below where they are today. Massive ice sheets scoured the land and revamped the ecological face of Earth, which was five degrees C cooler on average than it is currently.

The precise causes of the longer intervals between warm and cold periods are not yet sorted out. Volcanic eruptions may have played a significant role, as shown by the effect of El Chichón in Mexico and Mount Pinatubo in the Philippines. Tectonic events, such as the development of the Himalayas, may have influenced world climate. Even the impact of comets can influence short-term climatic trends with catastrophic consequences for life [see "What Caused the Mass Extinction? An Extraterrestrial Impact," by Walter Alvarez and Frank Asaro; and "What Caused the Mass Extinction? A Volcanic Eruption," by Vincent E. Courtillot; SCIENTIFIC AMERICAN, October 1990]. It is remarkable that despite violent, episodic perturbations, the climate has been buffered enough to sustain life for 3.5 billion years.

One of the most pivotal climatic discoveries of the past 30 years has come

OXYGEN RISING, CO_2 FALLING

ATMOSPHERIC COMPOSITION, shown by the relative concentration of various gases, has been greatly influenced by life on Earth. The early atmosphere had fairly high concentrations of water and carbon dioxide and, some experts believe, methane, ammonia and nitrogen. But levels of those gases have plummeted since then. After the emergence of living organisms, the oxygen that is so vital to our survival became more plentiful. Today carbon dioxide, methane and water exist only in trace amounts in the atmosphere.



IAN WORPOLE

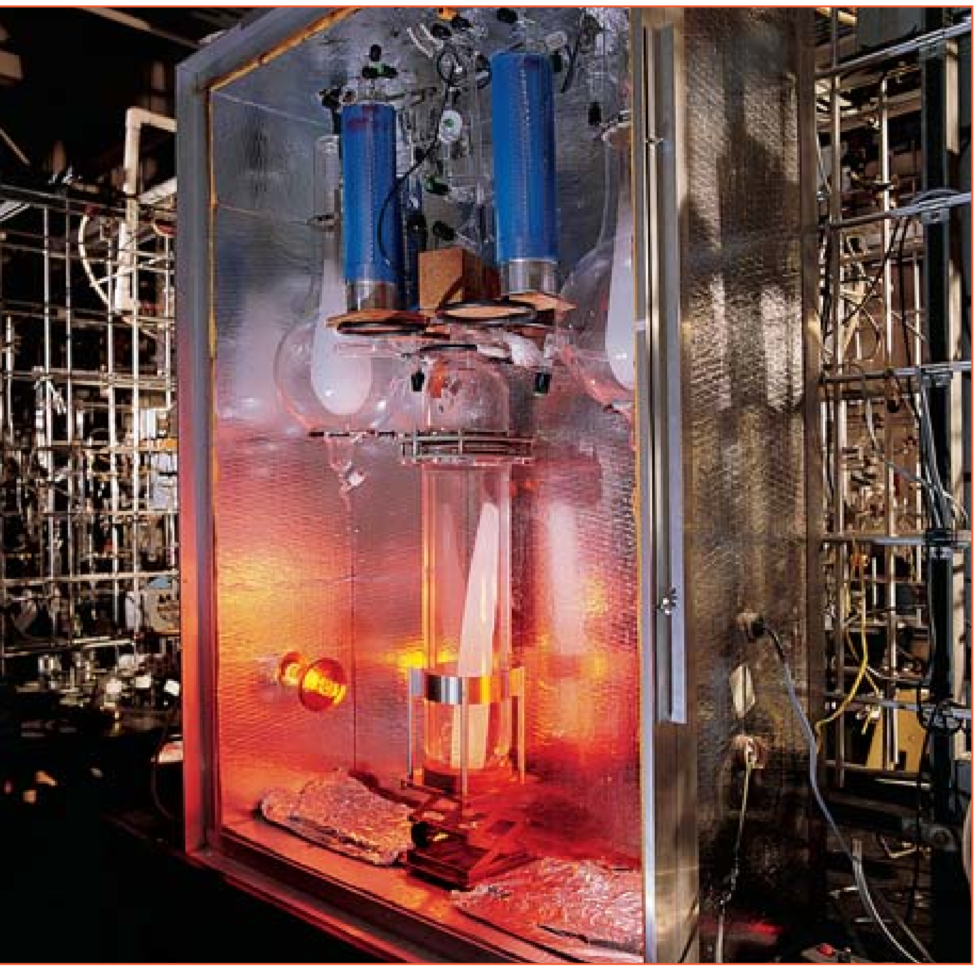
from ice cores in Greenland and Antarctica. When snow falls on these frozen continents, the air between the snow grains is trapped as bubbles. The snow is gradually compressed into ice, along with its captured gases. Some of these records can go back more than 500,000 years; scientists can analyze the chemical content of ice and bubbles from sections of ice that lie as deep as 3,600 meters (2.2 miles) below the surface.

The ice-core borers have determined that the air breathed by ancient Egyptians and Anasazi Indians was very similar to that which we inhale today—except for a host of air pollutants introduced over the past 100 or 200 years. Principal among these added gases, or pollutants, are extra carbon dioxide and methane. Since about 1860—the expansion of the Industrial Revolution—carbon dioxide levels in the atmosphere have increased more than 30 percent as a result of industrialization and deforestation; methane levels have more than doubled because of agriculture, land use and energy production. The ability of increased amounts of these gases to trap heat is what drives concerns about climate change in the 21st century [see "The Changing Climate," by Stephen H. Schneider; Scientific American, September 1989].

The ice cores have shown that sustained natural rates of worldwide temperature change are typically about one degree C per millennium. These shifts are still significant enough to have radically altered where species live and to have potentially contributed to the extinction of such charismatic megafauna as mammoths and saber-toothed tigers. But a most extraordinary story from the ice cores is not the relative stability of the climate during the past 10,000 years. It appears that during the height of the last ice age 20,000 years ago there was 50 percent less carbon dioxide and less than half as much methane in the air than there has been during our epoch, the Holocene. This finding suggests a positive feedback between carbon dioxide, methane and climatic change.

The reasoning that supports the idea of this destabilizing feedback system goes as follows. When the world was colder, there was less concentration of greenhouse gases, and so less heat was trapped. As Earth warmed up, carbon dioxide and methane levels increased, accelerating the warming. If life had a hand in this story, it would have been to drive, rather than to oppose, climatic change. It appears increasingly likely that when humans became part of this cycle, they, too, helped to accelerate warming. Such warming has been especially pronounced since the mid-1800s because of greenhouse gas emissions from industrialization, land-use change and other phenomena. Once again, though, uncertainties remain.

Nevertheless, most scientists would agree that life could well be the principal factor in the positive feedback between climatic change and greenhouse gases. There was a rapid rise in average global surface temperature at the end of the 20th century [*see illustration on opposite page*]. Indeed, the period from the 1980s onward has been the warmest of the past 2,000 years. Nineteen of the 20 warmest years on record have occurred since 1980, and the 12 warmest have all occurred since 1990. The all-time record high year was 1998, and 2002 and 2003 were in second and third places, respectively. There is good reason to believe that the decade of the 1990s would have been even hotter had not Mount Pinatubo erupted: this volcano put enough dust into the high atmosphere to block some incident sunlight, causing global cooling of a few tenths of a degree for several years.



ICE CORES from Greenland or Antarctica have provided scientists with a swatch cut from Earth's atmospheric history. As snow is compressed into ice, air bubbles trapped between the flakes are preserved. By analyzing the gases in these tiny chambers, researchers can determine the composition of the atmosphere more than 500,000 years ago.

TIM FULLER

Could the warming of the past 140 years have occurred naturally? With ever increasing certainty, the answer is no.

The box at the right shows a remarkable study that attempted to push back the Northern Hemisphere's temperature record a full 1,000 years. Climatologist Michael Mann of the University of Virginia and his colleagues performed a complex statistical analysis involving some 112 different factors related to temperature, including tree rings, the extent of mountain glaciers, changes in coral reefs, sunspot activity and volcanism.

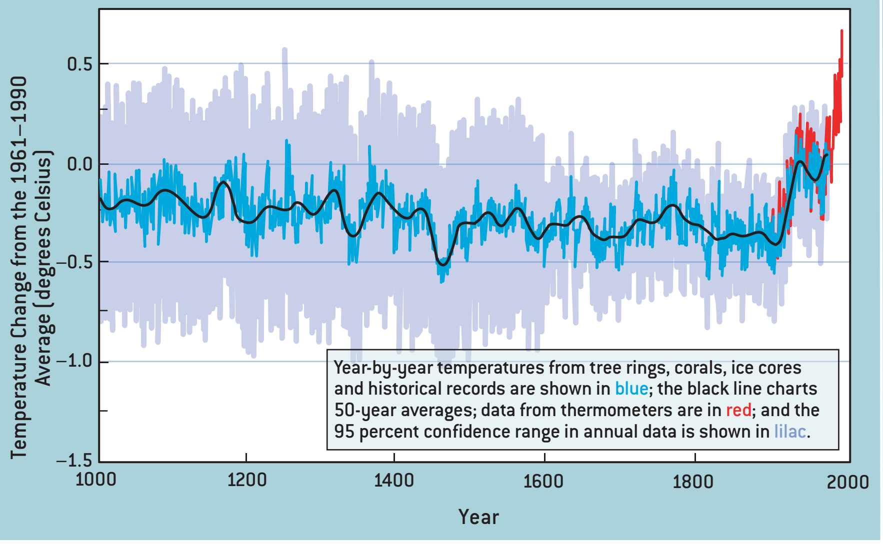
The resulting temperature record is a reconstruction of what might have been obtained had thermometer-based measurements been available. (Actual temperature measurements are used for the years after 1860.) As shown by the confidence range, there is considerable uncertainty in each year of this 1,000-year temperature reconstruction. But the overall trend is clear: a gradual temperature decrease over the first 900 years, followed by a sharp temperature upturn in the 20th century. This graph suggests that the decade of the 1990s was not only the warmest of the century but of the entire past millennium.

By studying the transition from the high carbon dioxide, low-oxygen atmosphere of the Archean to the era of great evolutionary progress about half a billion years ago, it becomes clear that life may have been a factor in the stabilization of climate. In another example—during the ice ages and interglacial cycles—life seems to have the opposite function: accelerating the change rather than diminishing it. This observation has led one of us (Schneider) to contend that climate and life coevolved rather than life serving solely as a negative feedback on climate.

If we humans consider ourselves part of life—that is, part of the natural system—then it could be argued that our collective impact on Earth means we may have a significant co-evolutionary role in the future of the planet. The current trends of population growth, the demands for increased standards of living and the use of technology and organizations to attain these growth-oriented goals all contribute to pollution. When the price of polluting is low and the atmosphere is used as a free sewer, carbon dioxide, methane, chlorofluorocarbons, nitrous oxides, sulfur oxides and other toxics can build up.

THE MODERN HEAT WAVE

DURING THE 20TH CENTURY, the rate of increase in temperatures and the duration of warming were much greater than for any of the preceding nine centuries. The 1990s were the warmest decade, and 1998 the warmest year, of the past millennium.



DRASTIC CHANGES AHEAD

IN THEIR REPORT *Climate Change 2001,* climate experts on the Intergovernmental Panel on Climate Change estimated that the world will warm between 1.4 and 5.8 degrees C by 2100. The mild end of that range—a warming rate of 1.4 degrees C per 100 years—is still 14 times faster than the one degree C per 1,000 years that historically has been the average rate of natural change on a global scale. Should the higher end of the range occur, then we could see rates of climatic change nearly 60 times faster than natural average conditions, which could lead to changes that many would consider dangerous. Change at this rate would almost certainly force many species to attempt to move their ranges, just as they did from the ice age/interglacial transition between 10,000 and 15,000 years ago. Not only would species have to respond to climatic change at rates 14 to 60 times faster, but few would have undisturbed, open migration routes as they did at the end of the ice age and the onset of the interglacial era. The negative effects of this significant warming—on health, agriculture, coastal geography and heritage sites, to name a few—could also be severe.

To make the critical projections of future climatic change needed to understand the fate of ecosystems on Earth, we must dig through land, sea and ice to learn as much from geologic, paleoclimatic and paleoecological records as we can. These records provide the backdrop against which to calibrate the crude instruments we must use to peer into a shadowy environmental future, a future increasingly influenced by us. SA

MORE TO EXPLORE

From Stone to Star: A View of Modern Geology. Claude J. Allègre. Harvard University Press, 1992.

Earth's Early Atmosphere. James F. Kasting in *Science,* Vol. 259, pages 920–926; February 12, 1993.

Scientists Debate Gaia: The Next Century. Edited by S. H. Schneider, J. R. Miller, E. Crist and P. J. Boston. MIT Press, 2004.

LUCY READING-IKKANDA

The North American continent may be more nomadic than most of its inhabitants

By Ian W. D. Dalziel

The little airplane banked to the right. From my seat on the port side I could see its shadow crossing the ice. The skis made it look rather like a duck coming in to land on water, webbed feet outstretched. As the pilot leveled the aircraft, a huge cliff came into view, the dark brown of its rocks contrasting sharply with the pristine whiteness of ice and snow that faded into the horizon.

The steeply inclined layers of this Precambrian sandstone were distorted by concertinalike folds. I took several photographs. As we rounded the cliff, another came into view. Resting on top of the sandstone was a thin capping of rock almost as white as the background: Cambrian limestone. “Fascinating,” I thought as I raised my camera again. “The basic geology here is very similar to that of western North America.”

My colleagues and I had come to the Pensacola Mountains of Antarctica to study how the two geologic subdivisions—East and West—of the icy continent relate to each other. East Antarctica is an old Precambrian shield lying to the south of Australia, India and Africa; West Antarctica is part of the geologically young and active volcanic “ring of fire” that surrounds the Pacific Ocean. The uplifted rim of the East Antarctic shield meets West Antarctica along the Transantarctic Mountains, of which the Pensacolas form a northern extension.

It had been a long trip down: 14 hours from Los Angeles to New Zealand in a commercial jet, 10 hours from New Zealand to McMurdo Station in Antarctica in a ski-equipped Hercules transport and, finally, five hours across the continent to the Pensacola Mountains, bypassing the South Pole en route. Now, after setting up our base camp, we were finally at the mountains near the southern margin of the same ocean that laps the beaches of Los Angeles.

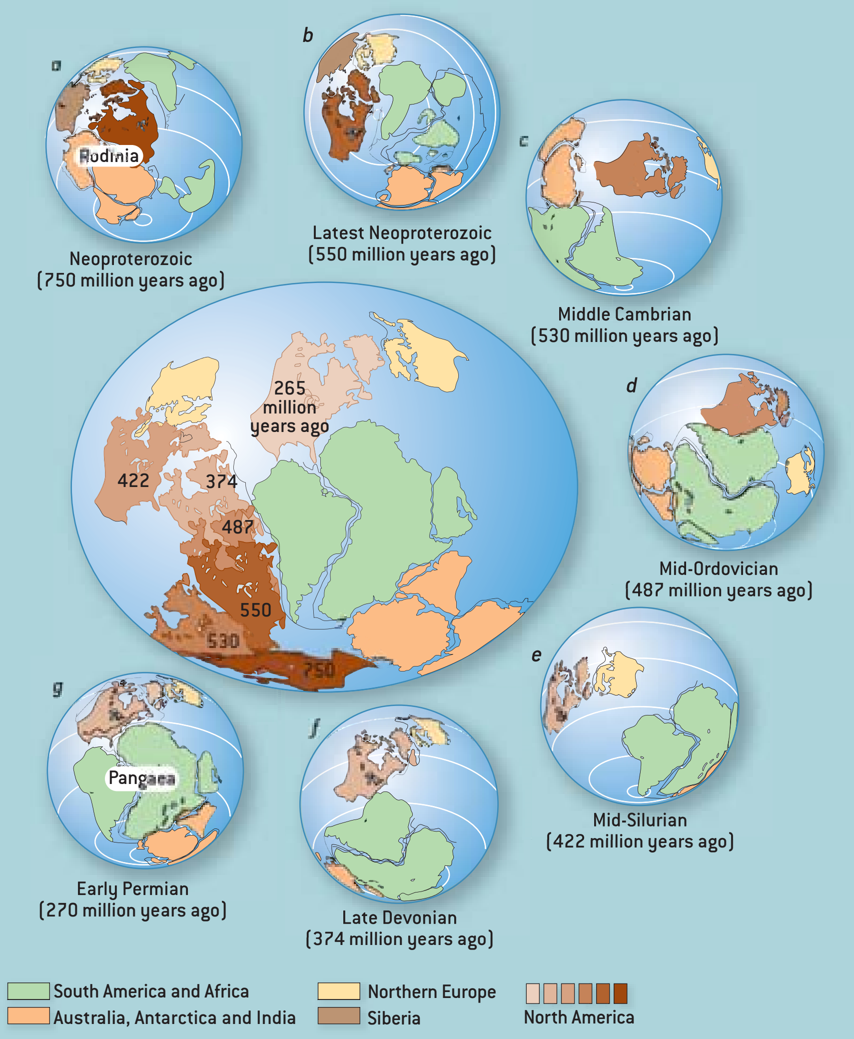
We still had to get to the rocks, however. In Antarctica such excursions take time. Having selected a possible crevasse-free landing site, our pilot brought the Twin Otter down for a “ski drag.” That is, he put some weight on the landing gear but maintained enough airspeed to take off again. We circled and carefully examined these tracks. Crevasses can be hidden under snow, but here there were no telltale signs of blue cracks. Coming around again, we touched down and stopped quickly so as to reduce the chance of hitting rough ice beneath the snow. It was a bumpy landing, nonetheless, although the aircraft appeared to have suffered only superficial damage. We roped ourselves together for safety and started to walk across the windblown snow to the base of the cliff, leaving our anxious pilot to examine the plane.

Earth

before Pangaea

NOMAD CONTINENT

NORTH AMERICA rifted apart from within the supercontinent Rodinia (*a*), traveled around neighboring continents and rejoined them to form Pangaea (*g*). On its way it collided with the South American shield (*d* and *f*) and northern Europe (*e*), leaving characteristic fragments. The central figure is a composite that juxtaposes the various positions of North America through its 500-million-year odyssey; the remaining continents are portrayed in their positions of 260 million years ago. The ancient latitudes of these landmasses are reconstructed using magnetic data.



FOSSIL CLUES

THE BOUNDARY between the two rock types exposed in the Pensacola Mountains is one of the most fundamental in Earth's history. After the birth of the planet 4.5 billion years ago came the four-billion-year-long interval of time known as the Precambrian. Toward the end of this era—about 750 million years ago, while the first soft-bodied, multicellular creatures were developing—the brown sandstones of the underlying Patuxent Formation we had just sighted were deposited. The strata were laid down in a rift valley that opened within the continental shield. As the rift deepened, rivers poured in, dropping their eroded soils onto the valley floor.

About 540 million years ago, an explosion of multicellular animal life ushered in the Cambrian period. Myriad cone-shaped skeletons of the creature Archaeocyatha collected in shallow seas that had advanced over the sandstone. These formed a reef along the rim of East Antarctica that was eventually transformed into limestone. (The cap on the Patuxent Formation is called the Nelson Limestone.) Because Archaeocyatha was a warm-water animal, what is now the western margin of the East Antarctic shield must have been situated in tropical latitudes during the Cambrian period.

The rifting event that led to the Patuxent sandstones' being deposited reflects the separation of East Antarctica from some other continental landmass. The divergence opened the Pacific Ocean basin about 750 million years ago. (Subsequently, igneous rocks from island volcanoes and material scraped off the subducting ocean floor accreted onto East Antarctica, forming West Antarctica.) This rifting occurred long before the supercontinent Pangaea—from which the present continents broke off—was formed. Pangaea was assembled only at the end of the Paleozoic era, approximately 250 million years ago. It started to fragment during the Jurassic period of the Mesozoic era, about 170 million years ago, creating the Atlantic and other young ocean basins.

Making our way up a ridge toward the top of the cliff, we saw that the lowest layers of the Cambrian strata—which lie below the limestone—were made of pink conglomerate and coarse sandstones. As the sea advanced over the deepening rift and the subsiding margin, it had ground the Precambrian rocks into boulders, pebbles and sand grains. The deposits became more fine-grained as we climbed, and the quartz sandstones immediately underneath the Nelson Limestone had the appearance of old friends. They were full of vertical worm burrows known as *Skolithus*.

These tubes are the only traces of ancient filter feeders, which extracted nutrients from sediments and left a clayey residue around their burrows. "Just like western North America," I noted out loud, "but then just like the Durness rocks of northwestern Scotland, too." Indeed, strata deposited by the seawater that advanced to cover most of the continents 540 million years ago—as shown by the presence of Cambrian seashores

GALEN ROWELL *Mountain Light* (*preceding pages*); IAN WORPOLE (*this page*) in such places as Wisconsin—are remarkably similar on all continents.

MATCHING MOUNTAINS...

THERE IS NOTHING like personal experience with rocks, however, to set a geologist thinking. My first impressions of the Transantarctic Mountains in 1987 raised a question that stayed near the forefront of my mind: Could the continent from which Antarctica rifted apart at the end of the Precambrian possibly have been western North America? Or were their margins at that distant time merely in similar environments on either side of an even more ancient Pacific Ocean basin?

The answer has far-reaching implications. The global paleogeography of the time ("paleo" is a prefix that geologists use to indicate "historical") is currently a mystery. To know how the continents were distributed could provide clues to the vast environmental alterations that preceded the Cambrian period. Late in Precambrian times there were several ice ages, and the oceanic and, presumably, atmospheric chemistry changed greatly. Multicellular animals evolved, heralding a biological profusion that included the far-distant ancestors of vertebrates and, hence, of human beings [see "End of the Proterozoic Eon," by Andrew H. Knoll; SCIENTIFIC AMERICAN, October 1991].

It is clearly difficult to map out with much certainty the geography of an ancient time on a dynamic planet with continents that move. Alfred Wegener and other pioneers of the theory of continental drift had noted that several North and South American mountain ranges truncated at the Atlantic margins match up neatly across the ocean with mountain ranges in Europe and Africa. Nowadays magnetic data and satellite images of the ocean floor showing fractures—appearing rather like railway tracks, along which the continents slid apart—allow us to reconstruct the supercontinent Pangaea very accurately.

A number of lines of evidence indicate that Pangaea was not the original configuration of the continents. When iron-bearing rocks solidify from lava, they become magnetized in the direction of Earth's magnetic field. The magnetization of rocks that congealed from pre-Mesozoic lava is quite different in North America and Africa, suggesting that in an earlier era these continents moved separately. Volcanic rocks that were fragments of ancient ocean floor have also been found in mountain ranges of Pangaea such as the Famatinian belt (Argentina), the Mozambique belt (Africa) and the older Appalachians. These early Paleozoic and Precambrian ophiolites—as the rocks are called—demonstrate that former ocean basins closed when the supercontinent amalgamated. Struck in the 1960s by the presence of early Paleozoic ophiolites in the Appalachian Mountains of the Maritime provinces in Canada, the imaginative Canadian geophysicist J. Tuzo Wilson asked: "Did the Atlantic Ocean open, close and then reopen?"

The quartz sandstones had the appearance of old friends.

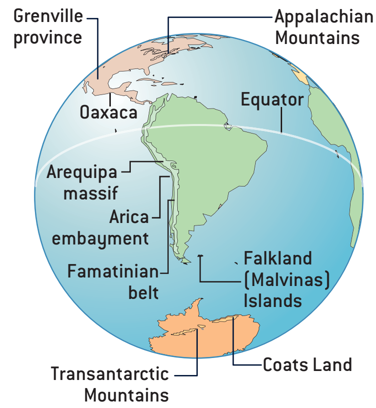
In reconstructing continental configurations prior to Pangaea, we get no help from the ocean floors. Although the Pacific Ocean basin already existed, ocean floor of such antiquity has long been thrust under the continents bordering the basin. Geologists therefore have no oceanic "railway map" for continental drift before Pangaea. We have to fall back on evidence from the continents themselves, just as Wegener did when attempting to reconstruct Pangaea before modern oceanography and satellites.

...AND MARGINS

WITHIN PANGAEA there are some ancient continental margins that have no obvious counterparts. The Pacific margins of North and South America, Antarctica and Australia were all formed near the end of the Precambrian, between 750 million and 550 million years ago. The Appalachian margin of Laurentia—the ancestral shield of North America—also rifted away from another continent at that time. Since Wilson asked his famous question, the counterpart to this margin has usually been assumed to have been western Europe and northwestern Africa. But there is no firm evidence for such a juxtaposition.

In 1989 I led another field trip to Antarctica, as part of the International Geological Congress hosted by the U.S. The object of the expedition was to help bring Antarctic geology—long the private domain of a very small group of especially hardy souls (even among geologists)—into the mainstream of global earth science. Various experts on the Himalayas, the European Alps, the Appalachians, the Rockies and many other regions participated.

Soon after, one of these scientists, Eldridge M. Moores, was browsing in



TRANSANTARCTIC MOUNTAINS mark an ancient boundary between East Antarctica and another continent, probably North America. The Dry Valleys (*see pages 14 and 15*) are cut into the uplifted margin of the chain. Features on today's Earth (*above*) record the travels of North America around other continents.

THE AUTHOR

IAN W. D. DALZIEL has been studying the geology of Antarctica, the Andes, the Caledonides and the Canadian Shield since earning his Ph.D. at the University of Edinburgh in 1963. Currently he is research professor and associate director at the Institute for Geophysics of the Jackson School of Geosciences at the University of Texas at Austin. In 1992 Dalziel received the Geological Society of London's Murchison Medal. In addition to his extensive geologic travels, he loves to visit wild places, preferably with his family. When in Austin, he sculls on Town Lake.

IAN WORPOLE

the library of the University of California at Davis when he came across a short article by Richard T. Bell and Charles W. Jefferson of the Geological Survey of Canada. They pointed out similarities between Precambrian strata in western Canada and eastern Australia and concluded that the Pacific margins of Canada and Australia might have been juxtaposed. Sensitized by his recent trip, Moores realized this would imply that the Pacific margins of the U.S. and Antarctica had been juxtaposed, a thought similar to my own. After some quick library research, he sent me a map highlighting the structural parallels in the interiors of the Laurentian and East Antarctic shields. "Is this crazy?" he asked.

Similarities in the internal structures of displaced continents can be powerful evidence of former juxtaposition. Moores drew particular attention to a report citing that along the Transantarctic Mountains—in a place called the Shackleton Range (after the famous British explorer Sir Ernest Shackleton)—lie rocks similar in age and character to those underneath much of New Mexico and Arizona. He also pointed out that roughly billion-year-old rocks like those characterizing the Grenville province—an aged band of rocks running along the eastern and southern margin of North America, from Labrador to Texas—had been found near one Antarctic shore. He called his hypothesis—the idea that the continents had been juxtaposed—SWEAT, for Southwest U.S.–East Antarctica.

Fired up by the possibility that my question might finally have an answer, I reproduced Moores's reconstruction using the PLATES software at our institute at the University of Texas at Austin. The program allows us to group together pieces of continents and move them over the globe with geometric precision. A short time later my colleague Lisa M. Gahagan and I had removed any uncertainties about matching the boundaries: the scale and general shape of the two old rifted margins were indeed compatible. Moreover, the boundary between the Grenville rocks of Texas and the older rocks of Arizona and New Mexico projected into Antarctica—just where I knew there was a similar boundary under the ice, between the Shackleton Range and some tiny rock outcrops along the frozen shores of the Weddell Sea. It seemed as if the rocks right under my feet, those that form the Llano uplift in Texas and from which the Texas State Capitol was built, were reappearing electronically in Antarctica!

If the western edge of North America was joined to East Antarctica and Australia, then some other continent must have rifted off the Appalachian margin. Paul F. Hoffman, now at Harvard University, and I have suggested that the eastern side of North America's Laurentian shield was wedged against the Precambrian shields of South America, known as Amazonia and Rio de la Plata. In manipulating the three shields on the computer screen, it occurred to me that the Labrador-Greenland prominence of Laurentia might have originated within the recess in the South American margin between Chile and southern Peru, often referred to as the Arica embayment. Both the promontory and the embayment are believed to date from late Precambrian times. But while they are of the same size and general shape, they were extensively modified when the Appalachian and Andean mountain chains rose. So a precise geometric fit is not to be expected.

Rocks under my feet were reappearing electronically in Antarctica!

CRYSTALLINE ENIGMA

MY SUGGESTION provides a possible explanation for a long-standing enigma of Andean geology. Along the otherwise youthful and active Peruvian margin are found 1.9-billion-year-old crystalline rocks. Hardolph A. Wasteneys, then at the Royal Ontario Museum, dated zircon crystals from the Arequipa massif, along the coast of southern Peru. He demonstrated that these rocks were highly metamorphosed when North America's Grenville Mountains were formed, 1.3 billion to 0.9 billion years ago. They may therefore represent a continuation of the Grenville province of eastern and southern North America into South America.

The hypothesis of a South American connection for the eastern margin of Laurentia unexpectedly brought my career full circle. I grew up in Scotland and cut my geologic teeth on its rocks. Northwestern Scotland and the submerged Rockall Plateau—off the western margin of the British Isles—remained part of North America until the North Atlantic Ocean basin had almost finished opening. Scotland was at the apex of the Labrador-Greenland promontory. When nestled (electronically) in the Arica embayment, the rocks of the Scottish Highlands that I studied for my doctoral degree in the 1960s appear to continue into equally old rocks of Peru and Bolivia. Given how well studied the Scottish Highlands are, they may provide critical tests for a former North America–South America connection.

Assuming the SWEAT hypothesis and the Pan-American connection, we can try to reconstruct the global distribution of continents and oceans in the late Precambrian. Most geologists believe that the relative areas occupied by continents and ocean basins have not changed since the late Precambrian. If, therefore, Antarctica, Australia, North America and fragments of South America were fused into a pre-Pangaean supercontinent, now named Rodinia, then there had to have been vast oceans elsewhere. Ophiolitic relics caught up within the continents indicate that these oceans lay between India and today's East Africa (the Mozambique Ocean) and within Africa and South America (the Pan-African and Braziliano oceans, respectively).

Between roughly 750 million and 550 million years ago these ocean basins were destroyed, and all the Precambrian nuclei of Africa, Australia, Antarctica, South America and India amalgamated into the supercontinent of Gondwana. It was during this time interval that the Pacific Ocean basin opened between Laurentia and the East Antarctic–Australian landmass. Isotopic dating of volcanic rocks in Newfoundland shows that the ocean basin between Laurentia and South America did not open until the beginning of the Cambrian. North America may therefore have separated out in a two-stage process.

Reconstructing the travels of North America requires an essential piece of information: the magnetization of ancient rocks. Such data allow geologists to figure out the latitude and orientation of the rocks when they formed. But because Earth's magnetic field is axially symmetrical, paleomagnetic measurements cannot tell us about the original longitude of the rocks. Present-day lava from Iceland and Hawaii, for example, could reveal to a geologist 100 million years from now the latitudes and the orientation of these islands but not their vast difference in longitude. It would not be apparent that the islands are in different oceans.

Traditional reconstructions of Laurentia always place its Appalachian margin opposite northwestern Africa during the Paleozoic era. I decided to plot the relation of North America to Gondwana differently, taking advantage of the fact that the longitude of the continent is not constrained by paleomagnetic data. It turned out that North America could have made what one of my graduate students referred to as an "end run" around South America during the Paleozoic, starting from next to Antarctica.

When Luis H. Dalla Salda, Carlos A. Cingolani and Ricardo Varela of the University of La Plata in Argentina saw the sketch of the end run, they became excited. They had recently proposed that a Paleozoic mountain belt, whose roots are exposed in the Andes of northern Argentina, could have formed when another continent collided with Gondwana. Moreover, the western margin of this Famatinian belt includes Cambrian and Lower Ordovician limestones (between 545 million and 490 million years old) containing trilobites characteristic of North America. Perhaps, they reasoned, this is a "geologic calling card" left behind when North America collided with South America during the Ordovician period, 450 million years ago.

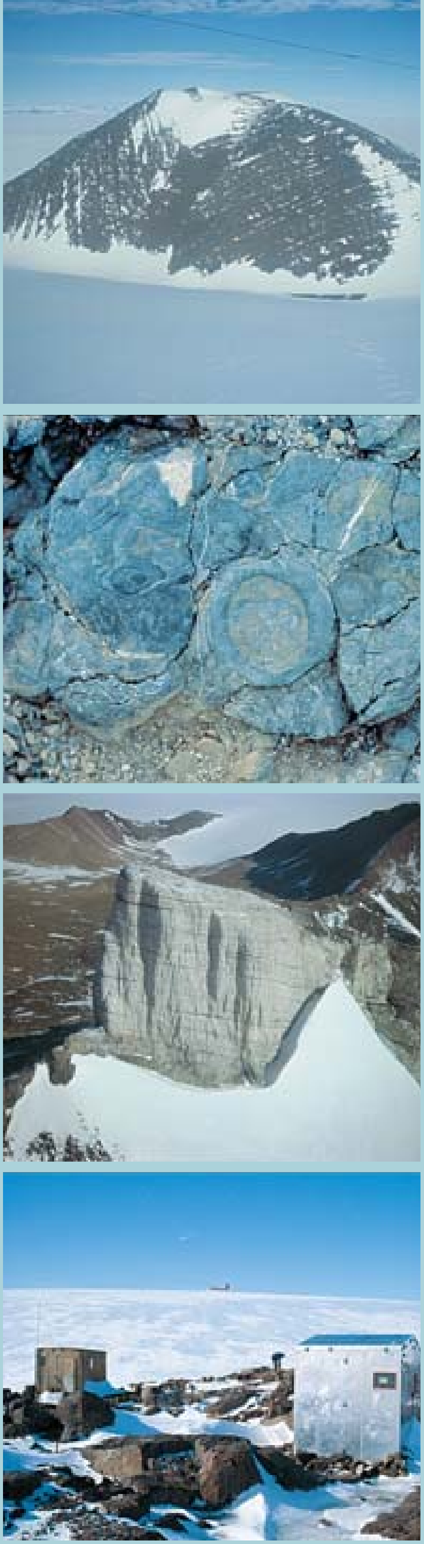
It appears that after rifting from South America at the end of the Precambrian, North America moved quite far away. During the Cambrian period, when Gondwana was undergoing glaciation, North America was equatorial. Ocean floor was then subducted beneath the South American craton, and North and South America collided again during the Ordovician. We think that the older part of the Appalachian Mountains, which terminates abruptly in Georgia, was once continuous with Argentina's Famatinian belt. This construction places Washington, D.C., close to Lima, Peru, during mid-Ordovician times.

END OF THE RUN

AFTER THE ORDOVICIAN collision, the continents separated again, apparently leaving North American limestone with its characteristic trilobites in northwestern Argentina. My Argentine colleagues and I have suggested that these rocks tore off the ancestral Gulf of Mexico, known as the Ouachita embayment. Blocks carried up by Andean volcanoes from below the limestones have recently been dated at around one billion years old, just like those of the Grenville province that probably occupied the embayment.

It is possible that the North and South American continents interacted again before North America finally collided with northwestern Africa to complete Pangaea. French geologists studying the Paleozoic sedimentary rocks of the Peruvian Andes have found that they are made of debris that must have eroded from a neighboring landmass. They assumed this continent, occupying the area now covered by the Pacific Ocean, to have been an extension of the Arequipa massif in Peru.

It may, however, have been North America. As Heinrich Bahlburg of the

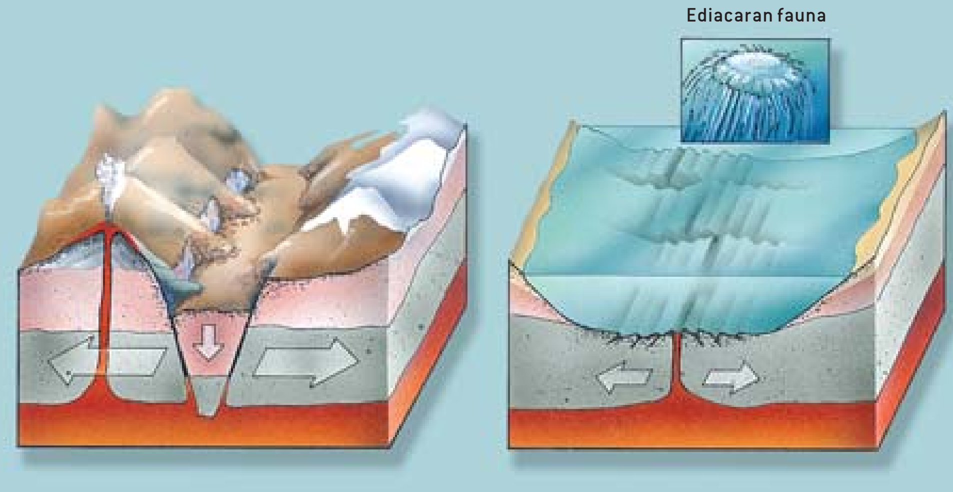


IAN DALZIEL (*top, bottom middle and bottom*); MARGARET REESE (*top middle*)

ROCK STRUCTURES in Antarctica provide clues to North America's voyages. Concertinalike folds (*top*) in the Patuxent Formation mark the Precambrian boundary between North America and Antarctica. As the two continents rifted apart, Archaeocyatha (*top middle*), among the first creatures with skeletons, formed a reef that fossilized into the Nelson Limestone cliff (*bottom middle*). Outcroppings of rhyolite lavas in Littlewood Nunataks, Coats Land (*bottom*), yield magnetic data that are being used to test the juxtaposition of North America and Antarctica. (The metal shack is a storm refuge for scientists from a nearby Argentine base.)

EARTH AND LIFE CO-EVOLVE

CONTINENTAL DRIFT is reflected in geology characteristic of each era. As Rodinia fragmented (*a*), sandstones of the Patuxent Formation in Antarctica were deposited. Glaciers lined the uplifted shoulders of the rift. In late Precambrian times (*b*), the Pacific Ocean opened up by spreading of the seafloor. Soft-bodied fauna simultaneously developed. Oceans advanced over the continents in Cambrian times (*c*), inhabited by hard-bodied trilobites and reef-forming Archaeocyatha. By the time tetrapods roamed, the continents had amalgamated into Pangaea (*d*). The Atlantic and Indian oceans opened as Pangaea broke apart (*e*) into today's world.



a 750 MILLION YEARS AGO

b 750–550 MILLION YEARS AGO

JANA BRENNING

University of Heidelberg in Germany has pointed out, ancient warm-water North American fauna mingle with cold-water fauna of southern Africa and the Falkland (Malvinas) Islands in the 400-million-year-old (Devonian) strata of northwestern South America. Together with a deformation along the eastern seaboard of North America known as the Acadian orogeny, and the truncation of mountain structures along the South American margin, the evidence points to Laurentia's sideswiping northwestern South America during the Devonian. There are even Ordovician limestones with South American trilobites—another calling card—at Oaxaca in Mexico. Only after North America finally moved away from the proto-Andean margin did the Andean Cordillera of the present day begin to develop.

Some 150 million years later North America returned to collide with northern Europe, Asia and Gondwana. Pangaea—with the Urals, the Armorican Mountains in Belgium and northern France, the Ouachitas and the youngest Appalachians as sutures—arose from the collisions of these continents. After a 500-million-year odyssey, North America had finally found a resting place. But not for long. In another 75 million years it separated from Africa as Pangaea broke up, to move toward its current position.

During the southern summer of 1993–1994—six years after my first glimpse of the Pensacola Mountains and glimmerings of North America's odyssey—I returned to Antarctica. This time, with my colleague Mark A. Helper, two graduate students and two mountaineers, I explored the Shackleton Range and Coats Land near the Weddell Sea. According to my computer simulations, this is where North America's Grenville rocks had projected 750 million years ago. Antarctic geologists have long regarded these areas as anomalous.

At the end of our visit to Coats Land, we roped together, picked up our ice axes and climbed back to another small aircraft. Weighing down our packs—and the aircraft, which groaned into the air—were the rock samples we had gleaned that day. In the laboratories of my colleagues Wulf A. Gose and James N. Connelly, we sat down to analyze those rocks.

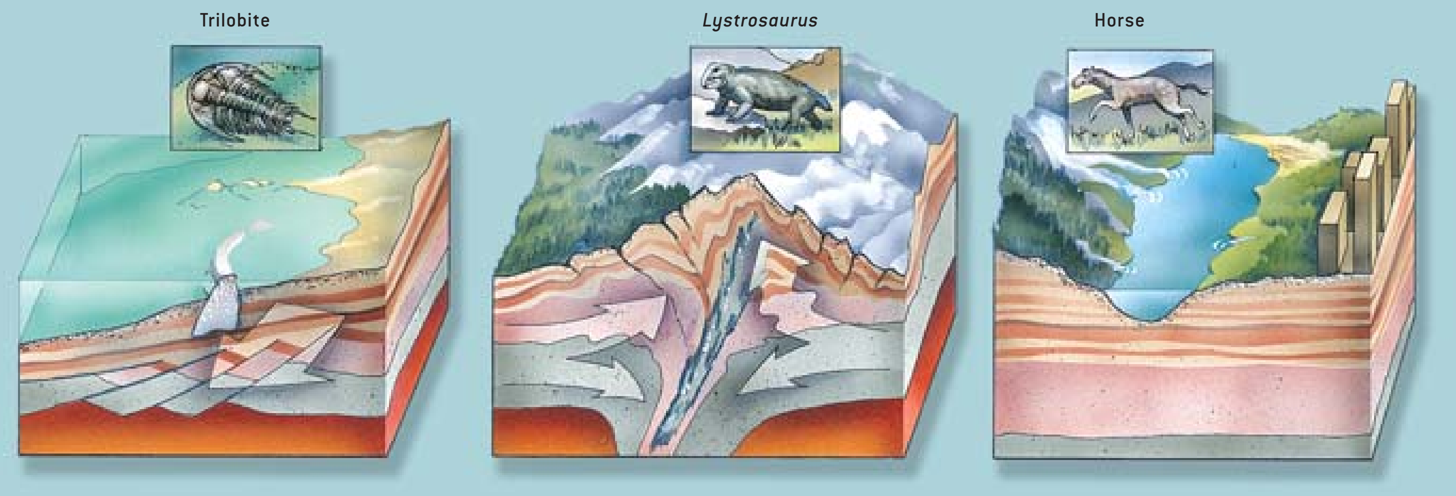
PERSUASIVE EVIDENCE

OUR IDEAS about how Earth looked before Pangaea, first described in this magazine in 1995, have stimulated a great deal of activity within the geologic community. They offered the first testable hypothesis regarding global geography in late Precambrian and early Paleozoic times—the critical era when single-celled organisms evolved into soft-bodied multicellular creatures, then invertebrates with hard shells and, ultimately, primitive vertebrates.

Over the past decade, interest in the Rodinian supercontinent that preceded Pangaea has spawned research centers and international programs to study this supercontinent's assembly, geography and fragmentation. One related result of this scientific ferment is the "snowball Earth" hypothesis, which proposes that Earth was covered with ice at sea level all the way to the equator 600 million to 700 million years ago, at the time of Rodinia's fragmentation and the formation of the Pacific Ocean basin.

The snowball Earth hypothesis posits an extreme global environment that challenges our understanding of climate past, present and future. If confirmed, it would mean that a dramatically chilly period directly preceded the explosion of multicellular life that occurred approximately 545 million years ago. Because forecasters rely on the distribution of continental landmasses in designing computer climate models, our rather esoteric study of ancient supercontinents has clearly taken on added significance in recent years.

With the absence of ocean floor predating Pangaea and the fragmentary nature of evidence from the continents, opinions regarding this period of Earth history inevitably differ. Some experts



c 540–500 MILLION YEARS AGO *d* 250 MILLION YEARS AGO *e* 180 MILLION YEARS AGO TO PRESENT

even doubt the very existence of the late Precambrian Rodinian supercontinent described in this article—doubts difficult to reconcile with the thousands of kilometers of preserved late Precambrian rifted continental margins.

Other researchers have used the same data we have relied on to reach radically different notions of the way this pre-Pangaean supercontinent may have looked. Instead of a connection between the southwestern U.S. and East Antarctica, for example, some experts propose that the U.S. Southwest and Mexico were connected to southeastern Australia. And an older idea has been revived, to the effect that Siberia was rifted from the proto-Pacific margin of North America. Nevertheless, two lines of evidence persuade me that our concept of the way Earth looked before Pangaea is the correct one.

First are the fruits of our 1993–1994 trip to Antarctica: the rock specimens we obtained from Coats Land. The paleomagnetic data obtained from those rocks do indeed show that this part of Antarctica could have been adjacent to the core of present-day North America when the rocks formed as volcanic deposits some 1.1 billion years ago. Extensive lava flows of this age lie exposed near Lake Superior and extend in the subsurface through Kansas to Trans-Pecos Texas, the Keeweenawan province. Although identical deposits exist throughout southern Africa's Umkondo province, my colleagues Jim Connelly here at Austin and Staci Loewy of the University of North Carolina at Chapel Hill have demonstrated that our Coats Land rocks contain lead isotopes that match those of North America's Keeweenawan province—but are quite distinct from the isotopic composition of the Umkondo lavas of Africa.

Second, evidence increasingly suggests that lower Paleozoic limestones of the Precordillera of northwestern Argentina originated in North America—yet another geologic calling card revealing North America's former presence off the Pacific margin of South America. Workers from both continents who have analyzed the rocks of Argentina's Precordillera have shown unequivocally that they originated in North America.

It remains unclear whether this ancient North American limestone arrived in South America as a Madagascar-like microcontinent or through transfer resulting from a continent-continent collision—as Italy was much later transferred from Africa to Europe when those two continents collided. Yet however they were transferred to South America, these limestone rocks offer the strongest possible evidence that North America did indeed make an end run around the Pacific margin of South America [*see illustration on page 17*] during Paleozoic times and that ancestral North America probably originated somewhere between the present Antarctic-Australian and South American–African parts of a pre-Pangaean supercontinent. SA

MORE TO EXPLORE

A Neoproterozoic Snowball Earth. Paul F. Hoffman, Alan J. Kaufman, Galen P. Halverson and Daniel P. Schrag in *Science,* Vol. 281, No. S381, pages 1342–1346; August 28, 1998.

Origin and Evolution of the Precordillera Terrane of Western Argentina: A Drifted Laurentian Orphan. Ricardo A. Astini and William A. Thomas in *Laurentia-Gondwana Connections before Pangea.* Special Paper of the Geological Society of America, No. 336, pages 1–20; 1999.

New Look at the Siberian Connection: No SWEAT. James W. Sears and Raymond A. Price in *Geology,* Vol. 28, No. 5, pages 423–426; May 2000.

Himalayan-Type Indenter-Escape Tectonics Model for the Southern Part of the Late Neoproterozoic–Early Paleozoic East African–Antarctic Origin. Joachim Jacobs and Robert J. Thomas in *Geology,* Vol. 32, No. 8, pages 721–724; August 2004.

***Salterella* in the Argentine Precordillera: An Early Cambrian Palaeobiogeographic Indicator of Laurentian Affinity.** Ricardo A. Astini, William A. Thomas and Ellis L. Yochelson in *Palaeogeography, Palaeoclimatology, Palaeoecology,* Vol. 213, Nos. 1–2, pages 125–132; October 7, 2004.

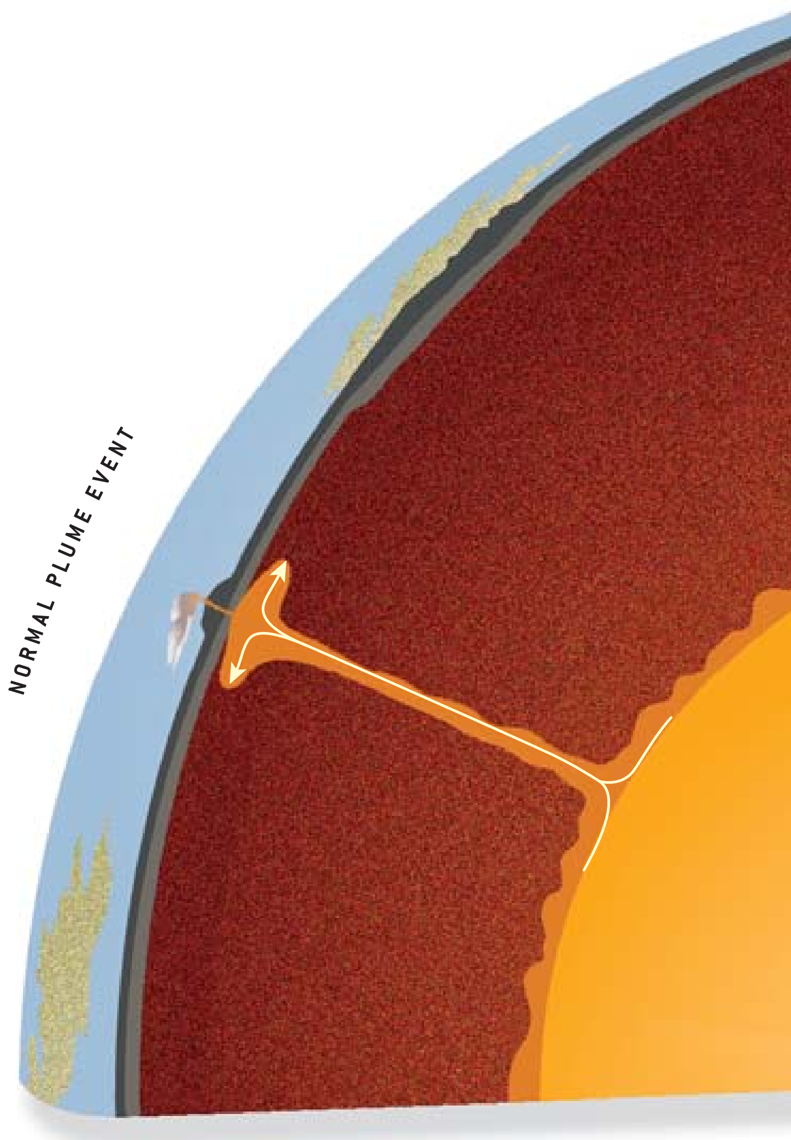
By Roger L. Larson

The Mid-Cretaceous

Earth has an erratic "heartbeat" that can release vast amounts of heat from deep within the planet. The latest "pulse" occurred 120 million years ago

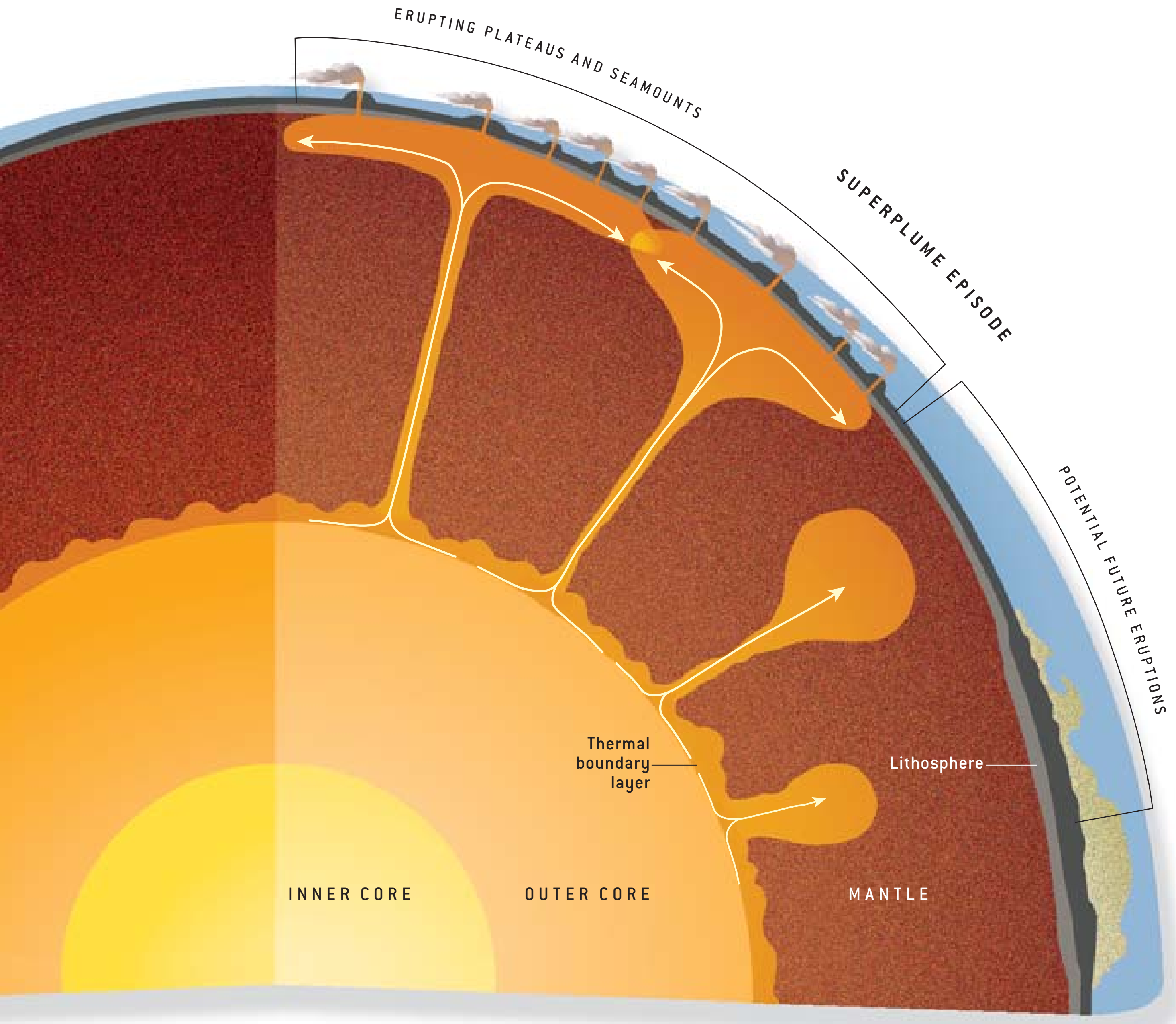
At one in the morning on December 13, 1989, I was awakened in my bunk onboard the scientific drillship *JOIDES Resolution* by sounds of celebration in the adjoining cabin. Because I had to relieve the watch at four anyway, I stumbled next door to join the party. The paleontologists in our expedition had just reported to my co–chief scientist, Yves Lancelot, now at the CNRS Center of Oceanology of Marseille in France, that microfossils of the Jurassic period had been recovered from the hole in the floor of the western Pacific Ocean that we were drilling more than three miles below us. Two days later the drill reached the volcanic basement—oceanic crust of Middle Jurassic age, about 165 million years old. A 20-year mystery was solved. At last, we had hard evidence of the world's oldest deep-sea sediments and volcanic rocks still in place from eons ago.

In succeeding days I reflected on why the quest had taken so long. My colleagues Clement G. Chase of the University of Arizona, Walter C. Pitman III of Lamont-Doherty Earth Observatory, Thomas W. C. Hilde of Texas A&M University and I had first considered the problem in the 1970s. The target was not a small one. We had predicted from geophysical data that an area in the western Pacific the size of the continental U.S. should be Jurassic in age, somewhere between 145 million and 200 million years old. But whenever we dredged or drilled in this area, we almost invariably recovered rocks called basalts, formed by volcanic eruptions during the mid-Cretaceous, generally ranging in age from 80 million to 120 million years but no older. The first such basalt samples were dredged from the Mid-Pacific Mountains in 1950 by an early expedition of the Scripps Institution of Oceanography. Until the *JOIDES* discovery, however, geologists



"SUPERPLUMES" of volcanic material build vast areas of oceanic plateaus and seamounts (*below*), compared with the small region affected by normal plumes (*opposite page*). The plumes are shown in a progressive sequence of events (*bottom to top*), which occurs for both superplumes and normal plumes: birth at the thermal boundary layer, ascent through the mantle, flattening at the base of the lithosphere and, finally, eruption at the surface.

Superplume Episode



had not made much progress in answering the questions concerning the origin of the seemingly ever present mid-Cretaceous basalts or the possible existence of underlying Jurassic material.

The 1989 discovery provided some qualitative answers. The older Jurassic sediments and oceanic crust were buried during the mid-Cretaceous epoch by what we now refer to as a "superplume" of volcanic material. Finally, our geophysical musings of the early 1970s could be supported with facts: the Jurassic existed in the western Pacific. We had samples of it locked away onboard the *JOIDES Resolution*.

The western Pacific's physiography is like a muddy New England road in March.

Because I am a geophysicist, I try to describe Earth and its processes quantitatively. I wanted to determine the size of the mid-Cretaceous superplume of the western Pacific, hoping to learn something of its origins. But saying that and doing it are two different things. What do you measure, and how do you measure it? I did not even know what "normal" was, so how could I describe the "anomalous" mid-Cretaceous superplume episode? The problem had to be expanded beyond the time and space framework of the mid-Cretaceous western Pacific. I decided to examine the rate of formation of oceanic crust—mainly volcanic rocks such as basalts that make up the solid basement underneath the seafloor—for all the ocean basins over their entire histories. Then the mid-Cretaceous anomaly, whatever it was, would stand out against the background. Clues to the timing of the next superplume might also appear.

At the time of the mid-Cretaceous, widespread volcanic eruptions covered or created vast amounts of ocean floor very quickly. Typically, though, seafloor spreading generates most of the oceanic crust in a slower, more regular way. In this process the crust becomes older symmetrically away from mid-ocean ridges where molten magma rises up out of Earth's mantle and then cools and solidifies. As new magma continues to rise, the older oceanic crust is rafted away from the eruption center and onto the flanks of the ridge. Thus, any particular parcel of crust is transported as if it were on one of two identical conveyor belts moving away from the mid-ocean ridge in opposite directions [see "The Mid-Ocean Ridge," by Kenneth C. Macdonald and Paul J. Fox; Scientific American, June 1990].

Areas of the ocean floor formed by spreading—known as abyssal plains—are covered with organized processions of abyssal hills and fracture zones running perpendicular to the mid-ocean ridges. Yet the western Pacific looks nothing like this. Its physiography is more like a muddy New England road in March. The seemingly randomly oriented chains of seamounts—undersea mountains that rise at least one kilometer from the ocean floor—and the oceanic plateaus that make up the "muddy road" of the western Pacific have no systematic age gradients across them. The only characteristic they share is that they are almost all from the mid-Cretaceous, to the extent that we even know their ages.

The first step in my investigation was to measure the changing rate of production of oceanic crust. To do this, I compiled information on the areas and ages of ocean floor and estimated the thickness of the crust. I was able to calculate this rate for the past 150 million years, nearly back to the maximum age of the world's ocean basins.

ONSET OF THE PULSE

THE BOX on the opposite page shows several of the geologic consequences of the mid-Cretaceous superplume event. The calculation of overall crustal production [*see third graph*] clearly shows the mid-Cretaceous superplume. This histogram of the world-total rate of crust production shows the plume's sudden onset 120 million to 125 million years ago, when formation of ocean crust nearly doubled in about five million years. Crustal production peaked soon after the onset of the pulse, tapering off more or less linearly over the next 70 million to 80 million years. Then, 30 million to 40 million years ago, it returned to values nearly the same as those before the episode. The mid-Cretaceous superplume in ocean crustal production stands out on a global scale. Yet the mere existence of the pulse does not indicate the reason for it.

I thought the key to the puzzle might lie in the development of oceanic plateaus and undersea mountain chains. During the mid-Cretaceous epoch, the rate of production of these formations jumped at the same time as the overall rate did, with a similar impulsive onset and a long succeeding taper back to normal values. Although the maximum amplitude at the height of this pulse was less than that for the world-total curve, the relative increase was much larger. Whereas total output of oceanic crust initially doubled, plateau and seamount production increased by a factor of five. So whatever produced the superplume episode also had the strongest effect on plateau and mountain-chain generation.

What causes these undersea plateaus and seamount chains? Independently, other investigators have converged on the notion that they result from plumes of material from deep

THE AUTHOR

ROGER L. LARSON first became acquainted with the oceans when he left Iowa State University with a bachelor's degree in geology. He headed west and five years later earned a Ph.D. in oceanography from the Scripps Institution of Oceanography at the University of California, San Diego. He became interested in the tectonic history of the western Pacific in 1971 as a research associate at the Lamont-Doherty Earth Observatory of Columbia University. That interest has continued during Larson's work as professor of oceanography at the University of Rhode Island. He has served as chief or co–chief scientist on a total of 15 oceanographic expeditions to the region.

IAN WORPOLE (*preceding pages*)

in Earth's mantle that have been overheated and thus rise buoyantly because of their reduced density [see "Large Igneous Provinces," by Millard F. Coffin and Olav Eldholm; Scientific American, October 1993]. In particular, oceanic plateaus result from the initial massive, rapid eruptions caused by these rising plumes. Such upwellings occasionally occur on the continents where we can study them directly. Exotically named regions such as the Paraná Basalts of Brazil, the Deccan Traps of western India and the Siberian Traps of northern Russia consist of vast fields of basalt flows, several hundred kilometers across and one or two kilometers thick [see "Volcanism at Rifts," by Robert S. White and Dan P. McKenzie; Scientific American, July 1989]. The oceanic plateaus are similar to their continental cousins, but they are even larger. For instance, the largest of the oceanic plateaus (the Ontong-Java Plateau of the western Pacific) is estimated to be 25 times bigger than the largest continental one (the Deccan Traps).

Seamount chains trail away from oceanic plateaus and result from material behind and below the head of the rising plume material. Because the plumes are relatively fixed, and the overlying tectonic plates drift horizontally on the surface of the planet, the subsequent eruptions in the mountain chains record the motions of the plates. Thus, these seamount chains should be oldest next to their parent oceanic plateau and trace a path of younger and younger seamounts that ends in an active volcano if the "chimney pipe" to the deep mantle is still alive. The best known of these seamount chains is the Hawaiian Islands, which extends underwater far to the northwest of the islands themselves. Its rising plume system exists today below the island of Hawaii, where volcanic eruptions continue to rumble. The islands and seamounts become successively older to the northwest as they are rafted away on the Pacific plate, which moves in a northwesterly direction over a fixed plume location.

Once I realized that the features of the ocean crust most affected by the mid-Cretaceous volcanic activity—seamount chains and plateaus—were each formed by plumes of mantle material, it was a small logical step to suppose that the entire anomaly resulted from plume activity on a much larger than normal scale. Because I live in a superlative-prone society, I named this a "superplume episode." The initial pulse of the superplume reached Earth's surface around 120 million years ago; the intense volcanic activity started suddenly and continued through the mid-Cretaceous, lasting tens of millions of years, gradually tapering off after that.

OVERHEATED PLUMES

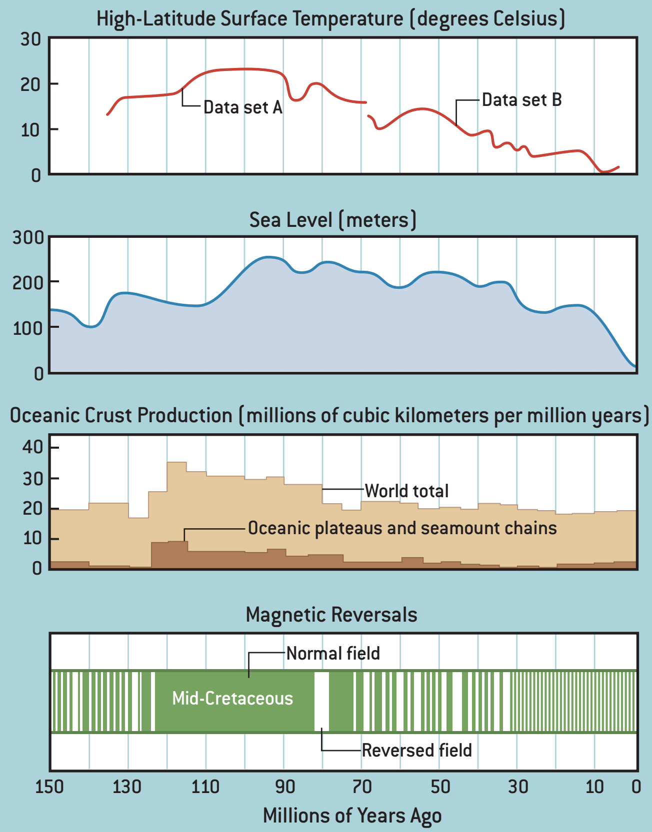
THE SUPERPLUME EPISODE was most likely caused by the upwelling of one or perhaps several enormous plumes that ascended through the easily deformed mantle, spread out at the base of Earth's more rigid outer shell, known as the lithosphere, and erupted onto the ocean floor. Although the Pacific was most strongly affected, evidence of the superplume event is also present in the Indian, South Atlantic and Caribbean oceans. The area of the Pacific involved may have been several thousand kilometers across, in sharp contrast with the size of regions affected by today's plume activity, which are usually one tenth that size in area.

I suspect the overheated plumes rise from the very base of the mantle and affect the process that causes reversals of Earth's magnetic field in the underlying outer core. There is a general inverse correlation between the production rate of crust formed by plumes and the frequency of reversals of Earth's magnetic field. For example, during periods of intense plume activity, including during the mid-Cretaceous, almost no magnetic reversals take place. Conversely, as is the case today, when plume activity is low, magnetic reversals occur at record pace.

How Earth's magnetic field actually reverses its polarity is a mystery [see "Probing the Geodynamo," on page 28]. Peter Olson of Johns Hopkins University and I think the correlation between crustal formation and magnetic field reversals may provide a clue to understanding how the reversals take place and to determining the source of the mantle plume

SUPERPLUME SEQUELS

GEOLOGIC CONSEQUENCES of the mid-Cretaceous superplume event include rising surface temperature and sea level. The superplume itself can be seen by the increase in the world-total rate of oceanic crust production; it is particularly evident in the rate of formation of oceanic plateaus and seamount chains. Additionally, reversals of Earth's magnetic field ceased during the superplume. At present, reversals occur regularly, indicating that plume activity is low.



IAN WORPOLE

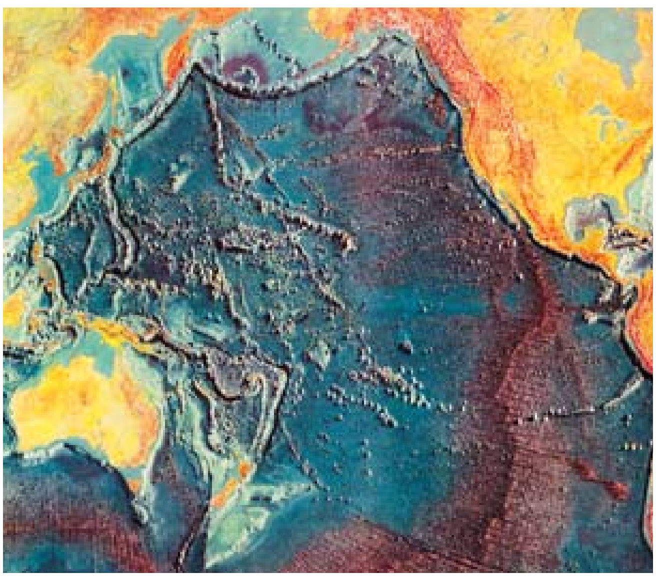
material. We believe an increase in the "boiling rate" of the core somehow causes magnetic reversals to become more infrequent. Additionally, the connection may reveal information about the advent of the next superplume.

Boiling iron within the outer core is the likeliest source of Earth's magnetic field. Such molten iron is an excellent electrical conductor, and the convective motion of the iron and its associated electrical field almost certainly generates Earth's magnetic field. The heat given off by the molten iron percolates through the core-mantle boundary, the lid to this boiling pot, by the process of conduction. The heat becomes trapped just above the boundary in the lowermost 100 to 200 kilometers of solid silicate rock of the mantle. This process continues until enough excess heat accumulates. Then the buoyancy of the overheated, less dense lower mantle overcomes the viscosity of the overlying cooler, more dense mantle rock. Huge plumes of mantle material rise nearly 3,000 kilometers through the mantle and eventually trigger volcanic eruptions at the surface. Ascending material removes heat from the lowermost mantle, allowing the outer core to boil even more vigorously than before.

In the mid-Cretaceous, my birthplace in Iowa was at the bottom of the ocean.

GLOBAL EFFECTS

THE MOST RECENT of these major overturns erupted 120 million to 125 million years ago as the mid-Cretaceous superplume episode. Much of the material that surfaced at this time left the muddy-road effect seen today on the western Pacific seafloor. An episode that nearly doubles the world-total rate of oceanic crustal production in a short period must have staggering geologic consequences. The mid-Cretaceous was characterized by several profound anomalies resulting from the superplume.



"MUDDY ROAD" of the western Pacific Ocean seafloor results from the intense volcanic activity of the mid-Cretaceous superplume event, which produced randomly oriented plateaus and undersea mountain chains. The ocean floor of the eastern Pacific, in contrast, shows the smooth, lineated physiography that is characteristic of crust formed by seafloor spreading.

First and probably least controversial is the rise in worldwide sea level to an elevation 250 meters or so higher than it is today. Assuming that the total amount of seawater in the planet's oceans is constant, a rise in the level of the sea surface is simply a reflection of a corresponding rise in the level of the seafloor. Ocean above newly formed crust is abnormally shallow because the crust and underlying lithosphere are still relatively warm, less dense and therefore expanded. As the two cool, they contract, allowing the seafloor to deepen. This phenomenon of expansion and contraction explains why oceanic ridges, where new crust is being formed, are raised above the older, deeper crust found on the flanks. If an abnormal amount of new crust is formed rapidly—as it was at the beginning of the mid-Cretaceous pulse—then the average seafloor level will be elevated, and the sea surface will rise accordingly. In the mid-Cretaceous, rising sea levels drowned much of what is dry land today; for example, my birthplace in Iowa was then at the bottom of the ocean. When the water receded, it left deposits of limestone and chalk, including the famous White Cliffs of Dover in England.

Earth's surface temperature also increased as a result of the superplume episode. When molten lava erupts, it releases certain chemicals, including carbon dioxide. Higher amounts of carbon dioxide in the mid-Cretaceous atmosphere led to a natural greenhouse effect that raised global temperatures by roughly 10 degrees Celsius. Studying the effects of elevated carbon dioxide levels during this period could reveal possible scenarios for Earth's climate in the future. Massive burning of fossil fuels and large-scale deforestation continue to increase the level of carbon dioxide in the modern atmosphere.

An excess amount of organic carbon and inorganic carbonate was also deposited during the mid-Cretaceous. The enhanced deposition is related to the elevations in sea level and air temperature, which, we have seen, resulted from the superplume episode. Tiny plants and animals, known as phytoplankton and zooplankton, make their living floating at shallow levels in the ocean where light can penetrate. Plankton apparently thrived during the mid-Cretaceous in the abnormally warm oceans that accompanied the natural warming of the atmosphere. Typically, when these organisms die, their bodies sink in the deep sea and quickly dissolve because of the extreme pressure of the overlying seawater. But during the mid-Cretaceous, many of the dead organisms fell instead on the drowned continents. The carbon from the exoskeletons did not dissolve in the shallow waters but was preserved. Some of it formed the White Cliffs, and some was buried more deeply and

KIMBERLITE DIAMONDS, from West Africa, formed more than a billion years ago and were transported to Earth's surface during the mid-Cretaceous superplume episode.

eventually turned to oil. The resulting oil constitutes up to 50 percent of the world's oil supply. Ironically, this outcome of the mid-Cretaceous greenhouse event may have created the fuel for the next greenhouse episode.

Other geologic anomalies associated with the mid-Cretaceous superplume include the placement of a very large percentage of Earth's diamond deposits. Diamonds are made of pure carbon atoms, squashed into the tightest, densest conceivable packing order by pressures that exist at least 200 to 300 kilometers below Earth's surface. Most diamonds are ancient even on geologic timescales, having formed more than a billion years ago, but according to Stephen E. Haggerty of the University of Massachusetts at Amherst, many of them were brought to the surface during the mid-Cretaceous. They were transported up volcanic structures called kimberlite diamond pipes (after a mining area in Kimberley, South Africa) that extend deep down into the crust and presumably into the upper mantle. The diamonds were probably torn loose from their sources within the mantle by rising plumes and brought up in their solid, original state.

The formation of most of the mountain ranges that edge the western coasts of North and South America was strongly controlled by the superplume episode. The Sierra Nevada Mountains of western North America and the Andes Mountains of western South America were created during the mid-Cretaceous by increased subduction of Pacific crust underneath western North and South America. Subduction occurs close to the continents when the oceanic lithosphere is thrust below the adjacent landmass and recycled into the mantle below. Remember that, because of the proximity of the erupting plumes, rates of seafloor spreading in the Pacific increased dramatically. What comes up must go down if Earth's diameter remains constant, so as production of ocean floor increased, so did subduction rates. Abnormally large amounts of oceanic crust were thrust deep under the western coastlines of North and South America. As the crust and accompanying oceanic sediments sank several hundred kilometers below Earth's surface, water within the sediments and crust lowered the melting temperatures of the solid material, causing it to become semiliquid as temperatures and pressures rose. Some of the adjacent continental crust also melted from frictional heating. This molten rock combination rose back to near the surface as its density lessened and then solidified to form the granite cores of the mountain ranges that are the spine of the western coast of the Americas.

SAM OGDEN

THE NEXT PULSE

OUR PLANET has clearly settled down from the effects of the mid-Cretaceous superplume event—the most recent superplume in Earth's history. But when the next superplume eruption will occur is a matter of speculation. We can get an idea of the possibility of modern superplumes by studying the time it takes earthquake waves to pass through the planet. This technique is similar to a CT scan used to obtain three-dimensional x-ray views of patients. Generally, if the earthquake waves arrive later than expected, then somewhere along their path they have passed through a part of Earth that is warmer than normal for that depth, and the warmer area has caused the waves to slow down. Conversely, if the waves arrive sooner than expected, then they have passed through a colder-than-normal area and speeded up. Warmer-than-normal areas in the deep mantle are often interpreted as mantle plumes.

Two large (more than 1,000 kilometers in all directions) warm anomalies in the deep mantle stand out, suggesting that superplumes are now forming deep within Earth. One of them is under French Polynesia in the South Pacific, the other under South Africa and the southeastern Atlantic. These have been called the South Pacific and African superplumes, although they lack the classic "mushroom cloud" appearance of a rising superplume. I suspect that the modern South Pacific superplume is the nearly exhausted remains of the mid-Cretaceous superplume, whereas the budding African superplume appears to be stalled in mid-ascent as a result of its chemically dense core.

I doubt that either of these present-day superplumes will erupt and reproduce the effects of the mid-Cretaceous, but I could be wrong—and the next superplume eruptions may be only a few million years away. I have no doubt that there will be more superplume eruptions in Earth's future, but no one can predict when the next one will occur. SA

MORE TO EXPLORE

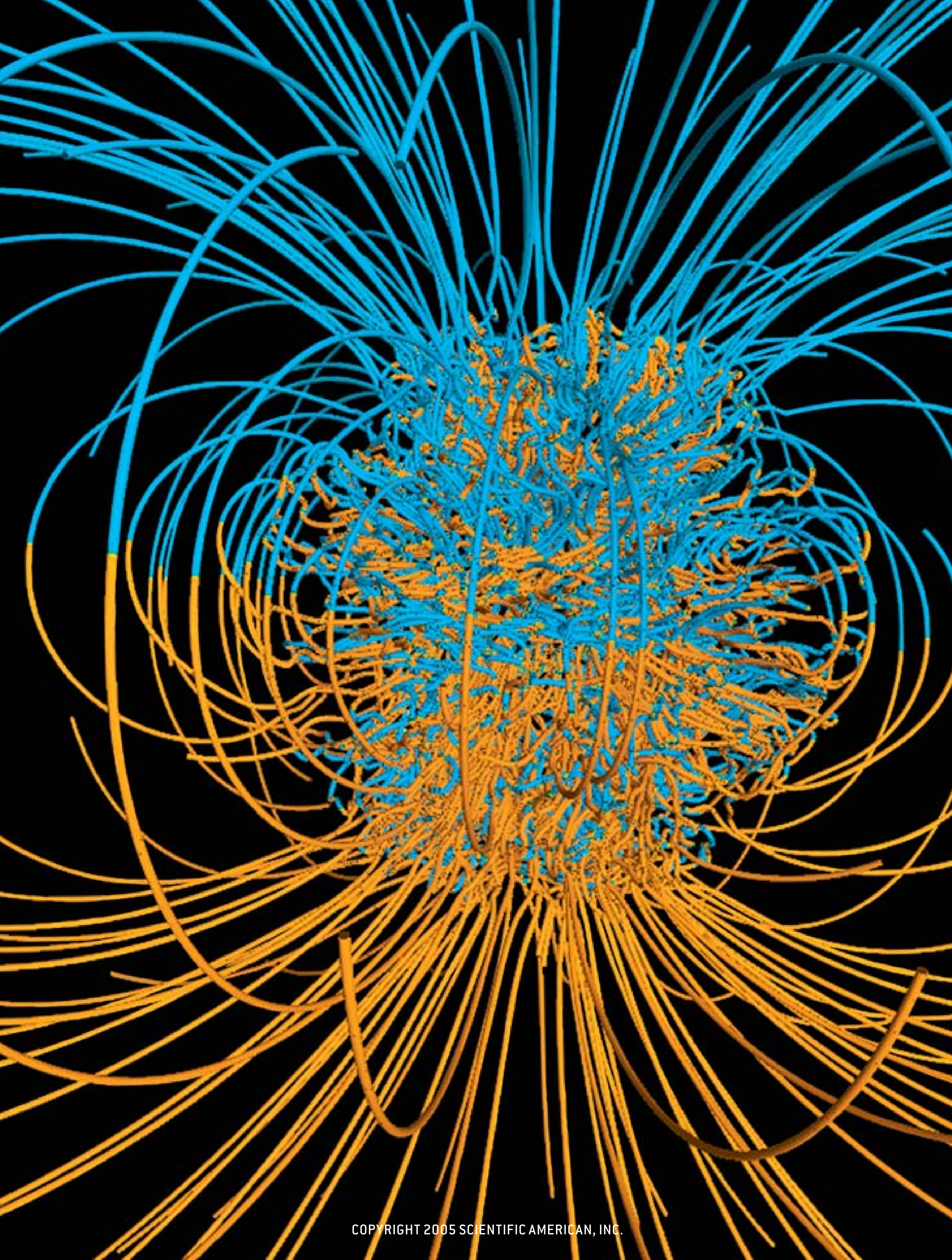
The Mid-Cretaceous Super Plume, Carbon Dioxide, and Global Warming. Ken Caldeira and Michael R. Rampino in *Geophysical Research Letters,* Vol. 18, No. 6, pages 987–990; June 1991.

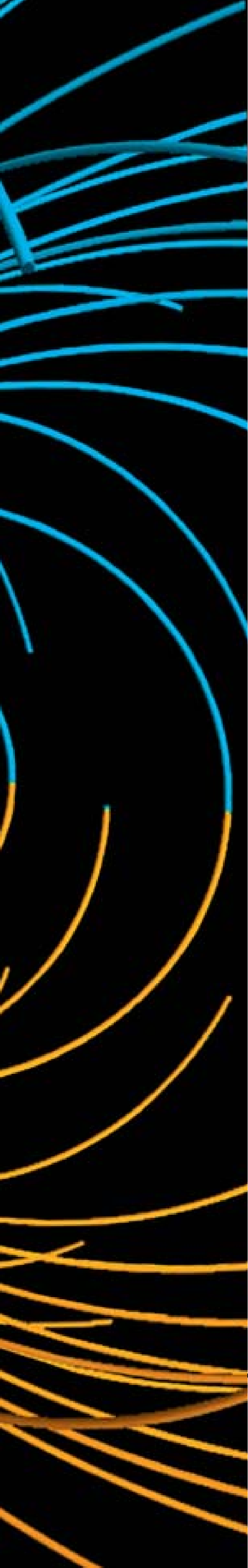
A Superplume in the Mantle. K. G. Cox in *Nature,* Vol. 352, No. 6336, pages 564–565; August 15, 1991.

Superplumes and Superchrons. Mike Fuller and Robin Weeks in *Nature,* Vol. 356, No. 6364, pages 16–17; March 5, 1992.

The Mesozoic Pacific: Geology, Tectonics, and Volcanism. Edited by Malcolm S. Pringle, William W. Sager, William V. Sliter and Seth Stein. American Geophysical Union, 1993.

Seismic Structure and the Origin of Hotspots and Mantle Plumes. Dapeng Zhao in *Earth and Planetary Science Letters,* Vol. 192, No. 3, pages 251–265; October 30, 2001.





Probing the Geodynamo

Scientists have wondered why the polarity of Earth's magnetic field occasionally reverses. Recent studies offer intriguing clues about how the next reversal may begin

By Gary A. Glatzmaier and Peter Olson

Most of us take it for granted that compasses point north. Sailors have relied on Earth's magnetic field to navigate for thousands of years. Birds and other magnetically sensitive animals have done so for considerably longer. Strangely enough, however, the planet's magnetic poles have not always been oriented as they are today.

Minerals that record past orientations of Earth's magnetic field reveal that it has flipped from north to south and back again hundreds of times during the planet's 4.5-billion-year history. But a switch has not occurred for 780,000 years—considerably longer than the average time between reversals, about 250,000 years. What is more, the primary geomagnetic field has lessened by nearly 10 percent since it was first measured in the 1830s. That is about 20 times faster than the field would decline naturally were it to lose its power source. Is this just a fluctuation in Earth's magnetic field, or could another reversal be on its way?

Geophysicists have long known that the source of the fluctuating magnetic field lies deep in the center of Earth. Our home planet, like several other bodies in the solar system, generates its own magnetic field through an internal dynamo. In principle, Earth's dynamo operates like the familiar electric generator, which creates electric and magnetic fields from the kinetic energy of its moving parts. In a generator, the moving parts are spinning coils of wire; in a planet or star, the motion occurs within an electrically conducting fluid. A vast sea of molten iron more than seven times the volume of the moon circulates at Earth's core, constituting the so-called geodynamo.

Until recently, scientists relied primarily on simple theories to explain the geodynamo and its magnetic mysteries. In the past 10 years, however, researchers have developed new ways to explore the detailed workings of the geodynamo. Satellites are providing clear snapshots of the geomagnetic field at Earth's surface, while new strategies for simulating Earth-like dynamos on supercomputers and creating physical models in the laboratory are elucidating those orbital observations. These efforts are providing an intriguing explanation for how polarity reversals occurred in the past and clues to how the next such event may begin.

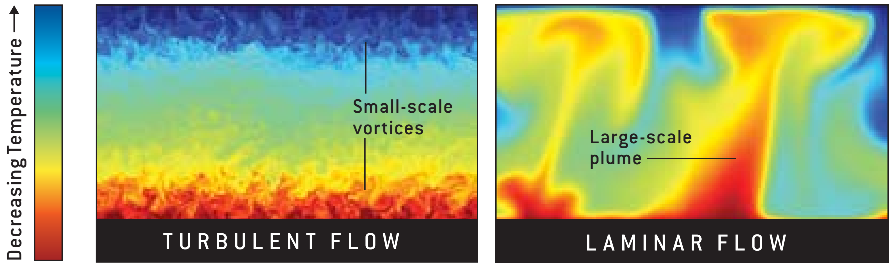
GARY A. GLATZMAIER

MAGNETIC LINES OF FORCE from a computer simulation of the geodynamo illustrate how Earth's magnetic field is much simpler outside the planet than inside the core (*tangled tubes at center*). At Earth's surface, the main part of the field exits (*long yellow tubes*) near the South Pole and enters (*long blue tubes*) near the North Pole.

DRIVING THE GEODYNAMO

BEFORE WE EXPLORE how the magnetic field reverses, it helps to consider what drives the geodynamo. By the 1940s physicists had recognized that three basic conditions are necessary for generating any planet's magnetic field. A large volume of electrically conducting fluid, the iron-rich liquid outer core of Earth, is the first of these conditions. This critical layer surrounds a solid inner core of nearly pure iron and underlies 2,900 kilometers of solid rock that form the massive mantle and the ultrathin crust of continents and ocean floors. The overlying burden of the crust and mantle creates average core pressures two million times greater than at the planet's surface. Core temperatures are similarly extreme—about 5,000 degrees Celsius, similar to the temperature at the surface of the sun.

These extreme environmental conditions set the stage for the second requirement of planetary dynamos: a supply of energy to move the fluid. The energy driving the geodynamo is part thermal and part chemical—both create buoyancy deep within the outer core. Like a pot of soup simmering on a burner, the core is hotter at the bottom than at the top. (The core's high temperatures are the result of heat that was trapped at the center of Earth during its formation.) That means the hot, buoyant iron in the lower part of the outer core tends to rise upward like blobs of hot soup. When the fluid reaches the top of the outer core, it loses some of its heat in the overlying mantle. The liquid iron then cools, becoming denser than the surrounding medium, and sinks. This process of transferring heat from bottom to top through rising and sinking fluid is called thermal convection—the second planetary condition needed for generating a magnetic field.



COMPLEX FLOW PATTERNS in Earth's molten outer core resemble two-dimensional computer simulations of turbulent convection (*left*). When running geodynamo simulations in three dimensions, however, scientists are limited to studying the larger plumes typical of laminar flow (*right*), which is akin to hot mineral oil rising through a lava lamp. Computers are so far incapable of resolving the much more complicated calculations associated with 3-D turbulent flow in Earth's core.

In the 1960s Stanislav Braginsky, now at the University of California at Los Angeles, suggested that heat escaping from the upper core also causes the solid inner core to grow larger, producing two extra sources of buoyancy to drive convection. As liquid iron solidifies into crystals onto the outside of the solid inner core, latent heat is released as a by-product. This heat contributes to thermal buoyancy. In addition, less dense chemical compounds, such as iron sulfide and iron oxide, are excluded from the inner core crystals and rise through the outer core, also enhancing convection.

For a self-sustaining magnetic field to materialize from a planet, a third factor is necessary: rotation. Earth's rotation, through the Coriolis effect, deflects rising fluids inside Earth's core the same way it twists ocean currents and tropical storms into the familiar spirals we see in weather satellite images. In the core, Coriolis forces deflect the upwelling fluid along corkscrewlike, or helical, paths.

That Earth has an iron-rich liquid outer core, sufficient energy to drive convection and a Coriolis force to twist the convecting fluid are primary reasons why the geodynamo has sustained itself for billions of years. But scientists need additional evidence to answer the puzzling questions about the magnetic field that emerges—and why it would change polarity over time.

MAGNETIC FIELD MAPS

A MAJOR DISCOVERY unfolded over the past five years as it became possible for scientists to compare accurate maps of the geomagnetic field taken 20 years apart. A satellite called Magsat measured the geomagnetic field above Earth's surface in 1980; a second satellite—Oersted—has been doing the same since 1999 [*see illustration on page 33*]. These satellite measurements provide an image of the magnetic field down to the level of the core-mantle boundary. But because of strong electric currents within the core, researchers cannot image the much more complicated and intense magnetic field inside the core, where the magnetic fluctuations originate. Despite the inherent limitations, several noteworthy observations came out of these efforts, including hints about the possible onset of a new polarity reversal.

Overview/*Turbulence Matters*

- The geologic record reveals that Earth's primary magnetic field switches polarity every so often, and researchers have long wondered why.
- Recent computer models of fluid motion in Earth's molten core have simulated an Earth-like magnetic field and associated polarity reversals. But because the fluid motion in these models is considerably simpler than the turbulent patterns thought to exist inside Earth, it is unclear how true to life these findings really are.
- Three-dimensional models capable of simulating turbulence, which are now under development, will one day resolve some of that uncertainty. In the meantime, satellite maps of the magnetic field and laboratory convection experiments are providing additional insight.

GARY A. GLATZMAIER

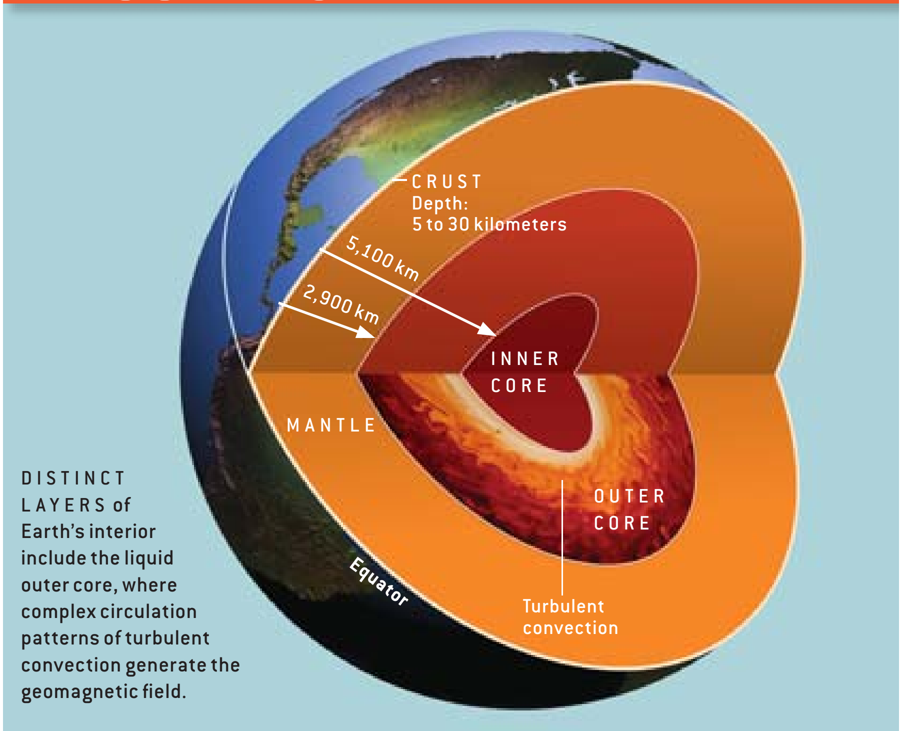
Although the geodynamo produces a very intense magnetic field, only about 1 percent of the field's magnetic energy extends outside the core. When measured at the surface, the dominant structure of this field is called the dipole, which most of the time is roughly aligned with Earth's axis of rotation. Like a simple bar magnet, this field's primary magnetic flux is directed out from the core in the Southern Hemisphere and down toward the core in the Northern Hemisphere. (Compass needles point to Earth's north geographic pole because the dipole's south magnetic pole lies near it.) But the satellite missions revealed that the flux is not distributed evenly across the globe. Instead most of the dipole field's overall intensity originates beneath North America, Siberia and the coast of Antarctica.

Ulrich R. Christensen of the Max Planck Institute for Solar System Research in Katlenburg-Lindau, Germany, suspects that these large patches come and go over thousands of years and stem from the ever evolving pattern of convection within the core. Might a similar phenomenon be the cause of dipole reversals? Evidence from the geologic record shows that past reversals occurred over relatively short periods, approximately 4,000 to 10,000 years. It would take the dipole about 100,000 years to disappear on its own if the geodynamo were to shut down. Such a quick transition implies that some kind of instability destroys the original polarity while generating the new polarity.

In the case of individual reversals, this mysterious instability is probably some kind of chaotic change in the structure of the flow that only occasionally succeeds in reversing the global dipole. Geoscientists who study the paleomagnetic record, such as Lisa Tauxe of the University of California at San Diego, find that dipole intensity fluctuations are common but that reversals are rare. The epochs between reversals vary in length from tens of thousands to tens of millions of years [*see illustration on page 35*].

Symptoms of a possible reversal-inducing change came to light when

A LOOK INSIDE



DISTINCT LAYERS of Earth's interior include the liquid outer core, where complex circulation patterns of turbulent convection generate the geomagnetic field.

JEN CHRISTIANSEN (*cutaway*); GARY A. GLATZMAIER (*outer core overlay*)

another group analyzed the Magsat and Oersted satellite maps. Gauthier Hulot and his colleagues at the Geophysical Institute in Paris noticed that sustained variations of the geomagnetic field come from places on the core-mantle boundary where the direction of the flux is *opposite* of what is normal for that hemisphere. The largest of these so-called reversed flux patches stretches from under the southern tip of Africa westward to the southern tip of South America. In this patch, the magnetic flux is inward, toward the core, whereas most of the flux in the Southern Hemisphere is outward.

PATCH PRODUCTION

ONE OF THE MOST significant conclusions that investigators drew by comparing the recent Oersted magnetic measurements with those from 1980 was that new reversed flux patches continue to form on the core-mantle boundary, under the east coast of North America and the Arctic, for example. What is more, the older patches have grown and moved slightly toward the poles. In the late 1980s David Gubbins of the University of Leeds in England—using cruder, older maps of the magnetic field—noticed that the proliferation, growth and poleward migration of these reversed flux patches account for the historical decline of the dipole.

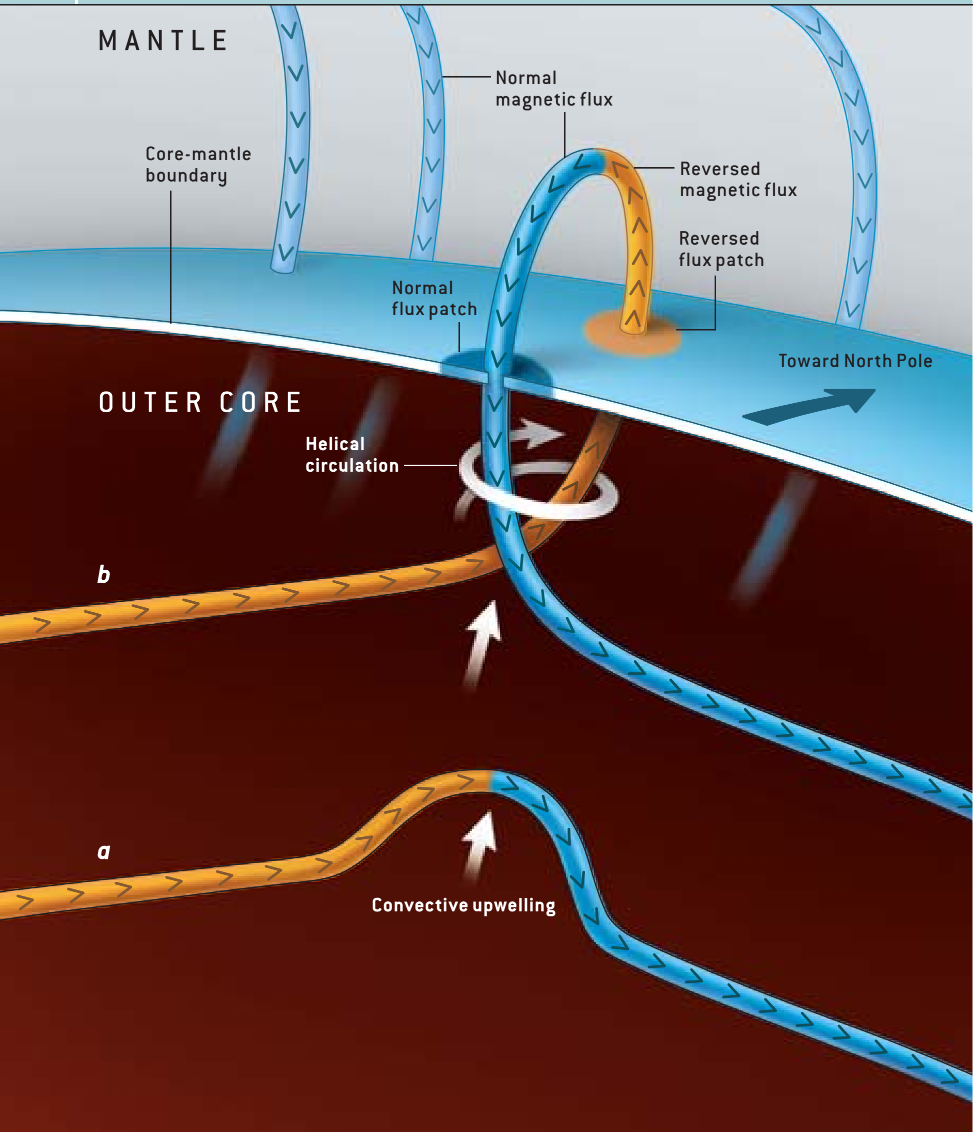
Such observations can be explained physically by using the concept of magnetic lines of force (in actuality, the field is continuous in space). We can think of these lines of force as being "frozen" in the fluid iron core so that they tend to follow its motion, like a filament of dye swirling in a glass of water when stirred. In Earth's core, because of the Coriolis effect, eddies and vortices in the fluid twist magnetic lines of force into bundles that look somewhat like piles of spaghetti. Each twist packs more lines of force into the core, thereby increasing the energy in the magnetic field. (If this process were to go on unchecked, the magnetic field would grow stronger indefinitely. But electrical resistance tends to diffuse and smooth out the twists in the magnetic field lines enough to suppress runaway growth of the magnetic field without killing the dynamo.)

Patches of intense magnetic flux, both normal and reversed, form on the core-mantle boundary when eddies and vortices interact with east-west-directed magnetic fields, described as toroidal, that are submerged within the core. These turbulent fluid motions can bend and twist the toroidal field lines into loops called poloidal fields, which have a north-south orientation. Sometimes the bending is caused by the rising fluid in an upwelling. If the upwelling is strong enough, the top of the poloidal loop is expelled from the core [*see box below*]. This expulsion creates a pair of flux patches where the ends of the loop cross the core-mantle boundary. One of these patches has normally directed flux (in the same direction as the overall dipole field in that hemisphere); the other has the opposite, or reversed, flux.

When the twist causes the reversed flux patch to lie closer to the geographic pole than the normal flux patch, the result is a weakening of the dipole, which is most sensitive to changes near its poles. Indeed, this describes the current situation with the reversed flux patch below the southern tip of Africa. For an actual planetwide polarity reversal to occur, such a reversed flux patch would grow and engulf the entire polar region; at the same time, a similar change in overall regional magnetic polarity would take place near the other geographic pole.

REVERSED FLUX PATCHES

REGIONS where the direction of magnetic flux is opposite that for the rest of the hemisphere arise when twisted magnetic fields occasionally burst above Earth's core. These reversed flux patches can weaken the main part of the magnetic field at Earth's surface, called the dipole, and may even signal the onset of a global polarity reversal. Reversed flux patches originate as fluid rising through the molten outer core pushes upward on roughly horizontal magnetic field lines within the core. This convective upwelling sometimes bends a line until it bulges (*a*). Earth's rotation simultaneously drives helical circulation of the molten fluid that can twist the bulge into a loop (*b*). When the upwelling force is strong enough to expel the loop from the core, a pair of flux patches forms on the core-mantle boundary.



SUPERCOMPUTER SIMULATIONS

TO FURTHER INVESTIGATE how reversed flux patches develop and how they may signal the onset of the next polarity reversal, researchers simulate the geodynamo on supercomputers and in laboratories. The modern era of computer dynamo simulations began in 1995, when three groups—Akira Kageyama, now of JAMSTEC, and his coworkers; Paul H. Roberts of U.C.L.A. and one of us (Glatzmaier); and Christopher A. Jones of the University of Exeter in England and his colleagues—independently developed numerical simulations that generated magnetic fields resembling the magnetic field at Earth's surface. Since then, simulations representing hundreds of thousands of years have demonstrated how convection can indeed produce patches of reversed magnetic flux on the core-mantle boundary—just like those seen in the satellite images. These patches come and go in computer simulations, but sometimes they lead to spontaneous magnetic dipole reversals.

Computer-generated polarity reversals provided researchers with the first rudimentary glimpse of how such switches may originate and progress [*see box on page 34*]. One three-dimensional simulation—which had to run for 12 hours a day every day for more than a year to simulate 300,000 years—depicted the onset of a reversal as a decrease in the intensity of the dipole field. Several patches of reversed magnetic flux, such as those now forming on the core-mantle boundary, then began to appear. But rather than extinguishing the magnetic field completely, the reversed flux patches created a weak field with a complex mix of polarities during the transition.

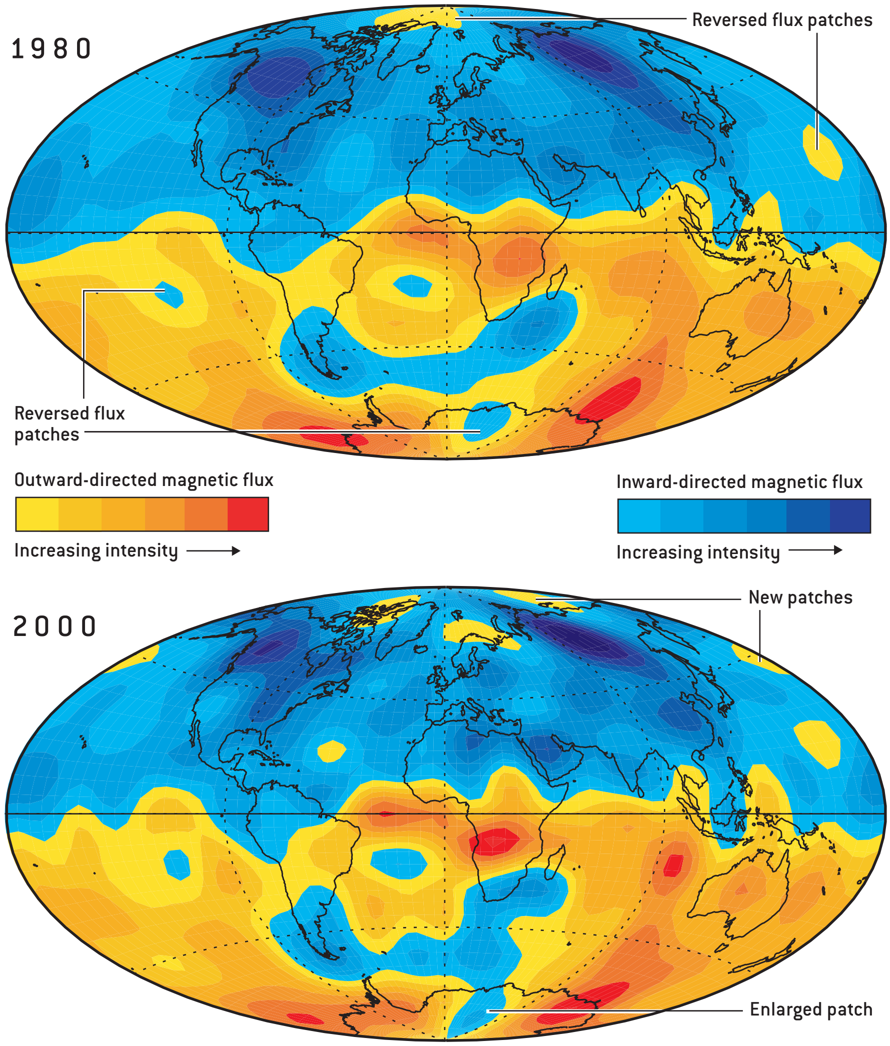
Viewed at the surface of the model Earth, the reversal of the dipole occurred when the reversed flux patches begin to dominate the original polarity on the core-mantle boundary. In total, it took about 9,000 years for the old polarity to dissipate and for the new polarity to take hold throughout the core.

WHAT MIGHT BE MISSING

BASED IN PART on these successes, computer dynamo models are proliferating rapidly. At last count, more than a dozen groups worldwide were using them to help understand magnetic fields that occur in objects throughout the solar system and beyond. But how well do the geodynamo models capture the dynamo as it actually exists in Earth? The truth is that no one knows for certain.

No computer dynamo model has yet simulated the broad spectrum of turbulence that exists in a planetary interior, primarily because massively parallel supercomputers are not yet fast enough to accurately simulate magnetic turbulence with realistic physical parameters in three dimensions. The smallest turbulent eddies and vortices in Earth's core that twist the magnetic field probably occur on a scale of meters to tens of meters, much less than what can be resolved with the current global geodynamo models on the current supercomputers. That means that all 3-D computer models of the geodynamo so far have simulated the simple, large-scale flow of laminar convection, akin to the hot mineral oil rising through a lava lamp.

To simulate the effects of turbulence in laminar models, investigators have used unrealistically large values for the fluid viscosity. To achieve realistic turbulence in a computer model, researchers must resort to a two-dimensional view. The trade-off is that 2-D flow cannot sustain a dynamo. These models do, however, suggest that the laminar flows seen in current geodynamo



PETER OLSON; SOURCE: GAUTHIER HULOT ET AL. IN *NATURE*, VOL. 416, PAGES 620–623; APRIL 11, 2002

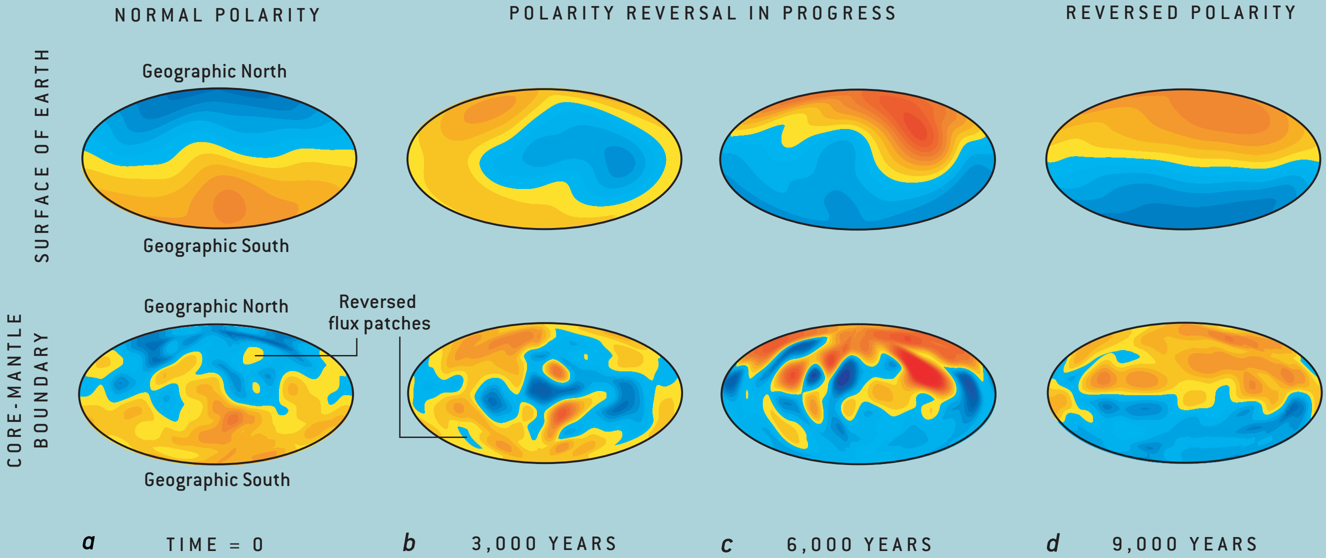
CONTOUR MAPS of Earth's magnetic field, extrapolated to the core-mantle boundary from satellite measurements, show that most of the magnetic flux is directed out from the core in the Southern Hemisphere and inward in the Northern Hemisphere. But in a few odd regions, the opposite is true. These so-called reversed flux patches proliferated and grew between 1980 and 2000; if they were to engulf both poles, a polarity reversal could ensue.

THE AUTHORS

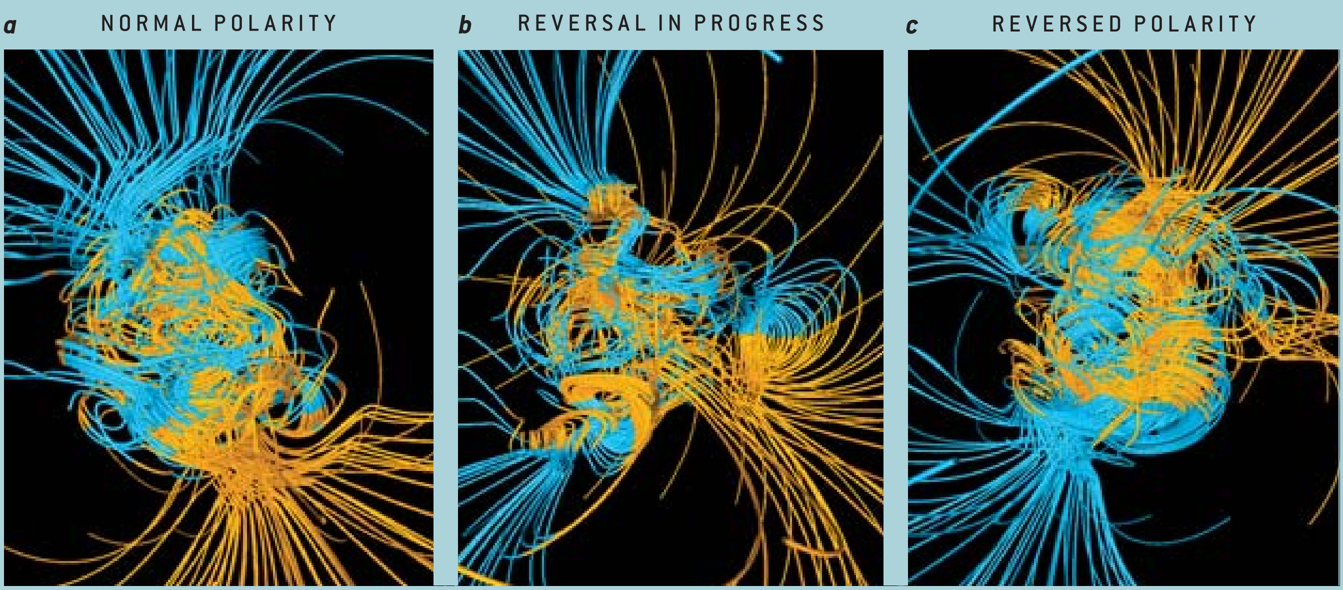
GARY A. GLATZMAIER and *PETER OLSON* develop computer models to study the structure and dynamics of the interiors of planets and stars. In the mid-1990s Glatzmaier, then at the Institute of Geophysics and Planetary Physics (IGPP) at Los Alamos National Laboratory, created (together with Paul H. Roberts of the University of California, Los Angeles) the first geodynamo simulation that produced a spontaneous magnetic dipole reversal. Glatzmaier has been a professor in the department of earth sciences and IGPP at the University of California, Santa Cruz, since 1998. Olson is particularly interested in how Earth's core and mantle interact to produce geomagnetic fields, plate tectonics and deep mantle plumes. He joined the department of earth and planetary sciences at Johns Hopkins University in 1978, where he has introduced geophysics to more than 1,000 students.

SIMULATED POLARITY REVERSALS

THREE-DIMENSIONAL COMPUTER simulations of the geodynamo are now capable of producing spontaneous reversals of the magnetic dipole, offering scientists a way to study the origin of reversals preserved in the paleomagnetic record [*see timeline on opposite page*]. One simulated switch from a model co-developed by one of us (Glatzmaier) occurred over a 9,000-year interval. This event is depicted as maps of the vertical part of the magnetic field at Earth's surface and at the core-mantle boundary, where the field is more complex. Models using magnetic field lines provide a third way to visualize a polarity reversal.

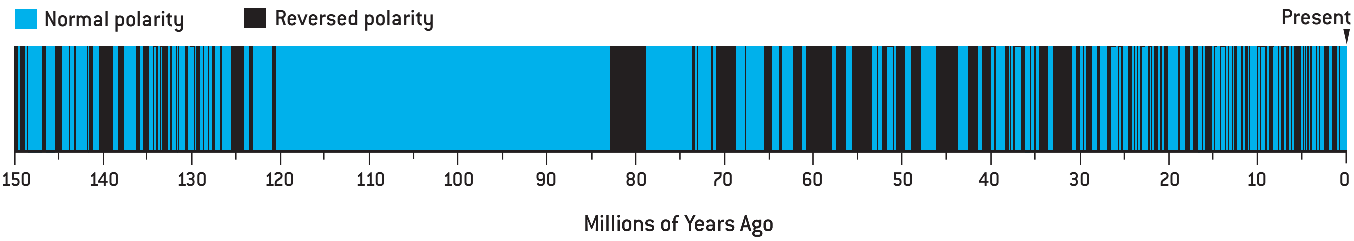


MAGNETIC FIELD MAPS start off with normal polarity, in which most of the overall magnetic flux points out from the core (*yellow*) in the Southern Hemisphere and in toward the core (*blue*) in the Northern Hemisphere (*a*). The onset of the reversal is marked by several areas of reversed magnetic flux (*blue in the south* and *yellow in the north*), reminiscent of the reversed flux patches now forming on Earth's core-mantle boundary. In about 3,000 years the reversed flux patches have decreased the intensity of the dipole field until it is replaced by a weaker but complex transition field at the core-mantle boundary (*b*). The reversal is in full swing by 6,000 years, when the reversed flux patches begin to dominate over the original polarity on the core-mantle boundary (*c*). If viewed only at the surface, the reversal appears complete by this time. But it takes an additional 3,000 years for the dipole to fully reverse throughout the core (*d*).



COMPUTER MODEL illustrates the magnetic field within the core (*tangled lines at center*) and the emerging dipole (*long curved lines*) 500 years before the middle of a magnetic dipole reversal (*a*), at the middle (*b*), and 500 years after that (*c*).

GARY A. GLATZMAIER; SIMULATED AT THE PITTSBURGH SUPERCOMPUTING CENTER



POLARITY REVERSALS occurred many times in the past 150 million years and probably long before that as well. Scientists discovered these reversals by studying volcanic rocks, which contain magnetic minerals. When such a rock cooled long ago, its magnetic field lined up in the direction of Earth's magnetic field at that location and time. The rock retains that ancient magnetic orientation indefinitely, unless it is heated to a temperature high enough to completely remagnetize it.

JEN CHRISTIANSEN; SOURCE: D. V. KENT AND W. LOWRIE IN *TIMESCALES OF THE PALEOMAGNETIC FIELD*. EDITED BY J.E.T. CHANNELL ET AL.; AGU, 2004

simulations are much smoother and simpler than the turbulent flows that most likely exist in Earth's core.

Probably the most significant difference is in the paths the fluid follows as it rises through the core. In simple laminar convection simulations, large plumes stretch all the way from the bottom of the core to the top. In the turbulent 2-D models, on the other hand, convection is marked by multiple small-scale plumes and vortices that detach near the upper and lower boundaries of the core and then interact within the main part of the convection zone.

Such differences in the patterns of fluid flow could have a huge influence on the structure of Earth's magnetic field and the time it takes various changes to occur. That is why investigators are diligently pursuing the next generation of 3-D models. Someday, maybe a decade from now, advances in computer processing speeds will make it possible to produce strongly turbulent dynamo simulations. Until then, we hope to learn more from laboratory dynamo experiments now under way.

LABORATORY DYNAMOS

A GOOD WAY to improve understanding of the geodynamo would be to compare computer dynamos (which lack turbulence) with laboratory dynamos (which lack convection). Scientists had first demonstrated the feasibility of lab-scale dynamos in the 1960s, but the road to success was long. The vast difference in size between a laboratory apparatus and the actual core of a planet was a vital factor. A self-sustaining fluid dynamo requires that a certain dimensionless parameter, called the magnetic Reynolds number, exceed a minimum numerical value, roughly 10.

Earth's core has a large magnetic Reynolds number, probably around 1,000, primarily because it has a large linear dimension (the radius of the core is about 3,485 kilometers). Simply put, it is very difficult to create a large magnetic Reynolds number in small volumes of fluid unless you can move the fluid at extremely high velocities.

The decades-old dream of generating a spontaneous magnetic field in a laboratory fluid dynamo was first realized in 2000, when two groups—one led by Agris Gailitis of the University of Latvia and one by Robert Stieglitz and Ulrich Müller of the Karlsruhe Research Center and Fritz Busse of the University of Bayreuth, both in Germany—independently achieved self-generation in large volumes of liquid sodium. (Liquid sodium was used because of its high electrical conductivity and low melting point.) Both groups found ways to achieve high-speed fluid flow in a system of one- to two-meter-long helical pipes, resulting in the critical magnetic Reynolds number of about 10.

These experimental results support the theory, which gives us a measure of confidence when we apply our theoretical ideas about dynamos to Earth and other planets. In labs across the world—at the University of Grenoble in France, the University of Maryland, the University of Wisconsin–Madison and the New Mexico Institute of Mining and Technology—scientists are developing the next generation of lab dynamos. To better simulate Earth-like geometry, these experiments will stir the liquid sodium inside massive spherical chambers—the largest nearly three meters in diameter.

Besides the ongoing plans for more realistic laboratory dynamos and 3-D computer simulations, the international satellite CHAMP (short for Challenging Minisatellite Payload) is charting the geomagnetic field with enough precision to directly measure its changes at the core-mantle boundary in real time. Investigators anticipate this satellite will provide a continuous image of the geomagnetic field over its five-year mission, allowing them to watch for continued growth of the reversed flux patches as well as other clues about how the dipole field is waning.

We expect that a synthesis of these three new approaches—satellite observations, computer simulations and lab experiments—will occur in the next decade or two. With a more complete picture of the extraordinary geodynamo, we will learn whether our current ideas about the magnetic field and its reversals are on the right track. SA

MORE TO EXPLORE

Numerical Modeling of the Geodynamo: Mechanisms of Field Generation and Equilibration. Peter Olson, Ulrich Christensen and Gary A. Glatzmaier in *Journal of Geophysical Research,* Vol. 104, No. B5, pages 10383–10404; 1999.

Earth's Core and the Geodynamo. Bruce A. Buffett in *Science,* Vol. 288, pages 2007–2012; June 16, 2000.

Geodynamo Simulations: How Realistic Are They? Gary A. Glatzmaier in *Annual Review of Earth and Planetary Sciences,* Vol. 30, pages 237–257; 2002.

Recent Geodynamo Simulations and Observations of the Geomagnetic Field. Masaru Kono and Paul H. Roberts in *Reviews of Geophysics,* Vol. 40, No. 4, page 1013; 2002.

The Core-Mantle Boundary

This interactive zone may be the most dynamic part of the planet, directly affecting Earth's rotation and magnetic field

By Raymond Jeanloz and Thorne Lay

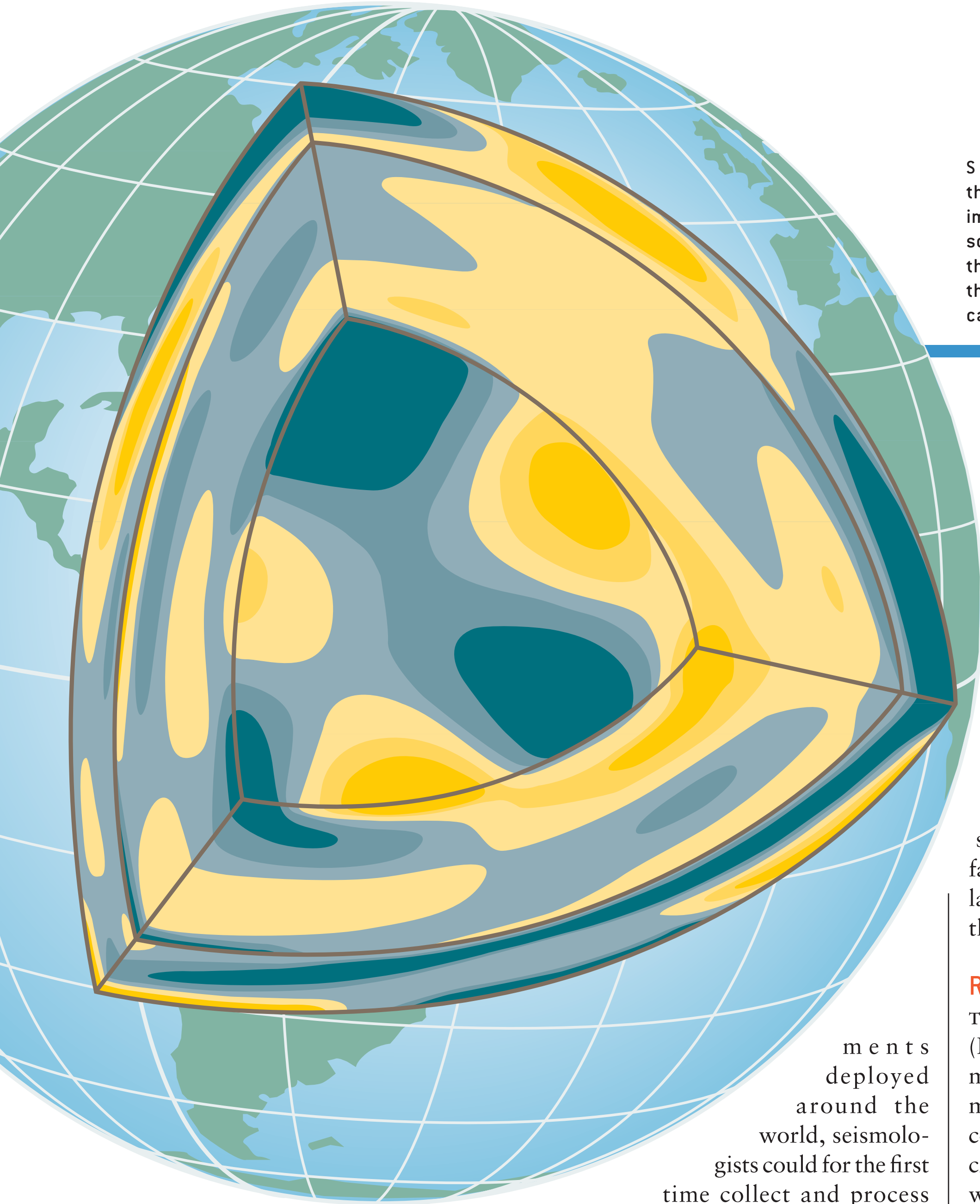
About 2,900 kilometers away—less than three days' drive, if that were possible—lies Earth's most dramatic structure. Largely ignored in past research, the remote region between the lowermost mantle and the upper core is proving to be crucial in understanding the chemical and thermal evolution of the planet. No longer regarded as simply a contact delineating the liquid-iron outer core from the rocky mantle, the core-mantle region may actually be the most geologically active zone of Earth. Its features seem to have changed immensely during Earth's history, and its physical properties vary from place to place near the bottom surface of the mantle. In fact, the physical changes across the interface between the core and mantle are more pronounced than those across the planetary surface separating air and rock.

The strong heterogeneity of the core-mantle boundary region is thought to influence many global-scale geologic processes [see "The Earth's Mantle," by D. P. McKenzie; Scientific American, September 1983]. The dynamics of the zone affect the slight wobbling of Earth's axis of rotation and characteristics of the geomagnetic field. Variations in the core-mantle region also modulate the convection in Earth's mantle, which is responsible for the movement of continents and tectonic plates.

The first hint that something unusual was going on at the depth where the core and mantle meet came in the mid-1930s. Vibrations generated by earthquakes provided the clue. Throughout most of the mantle, the speed of seismic waves increases as a function of depth. Furthermore, lateral variations in seismic-wave velocity are only minor. One can interpret these characteristics as meaning that Earth gets "simpler" with respect to depth—that is, the composition and structure of the planet become more uniform. In contrast, the great diversity of geologic structures and rocks observed underfoot reveal the surface to be the most complicated region.

Yet the velocity behavior of seismic waves holds only to a certain point. At the lowermost few hundred kilometers of the mantle, just before the core begins, the average speed of seismic waves does not increase appreciably, and more meaningful changes in velocity appear from region to region [*see box on page 38*]. The effect is subtle, amounting to only a few percent difference. Yet by geologic standards, these few percent represent enormous variations in structure, temperature or both. Early workers recognized the significance of the changes from the simple behavior in the overlying lower mantle and consequently named this region, which was deduced to be about 200 to 400 kilometers thick, the D″ layer.

The origin of the layer's name (pronounced "dee double prime") is more historic than poetic. Early geologists had labeled the parts of the deep Earth with letters of the alphabet, rather than as crust, mantle and core. This form of identification,



SEISMIC-WAVE VELOCITIES differ throughout Earth's interior, as depicted in this image generated by seismic tomography. In some regions the waves move more quickly than is average for that depth (*blues*); in others the waves are slower (*yellows*). Such variations can suggest differences in composition.

ADAM M. DZIEWONSKI *Harvard University* AND JOHN H. WOODHOUSE *University of Oxford*; PHOTOSHOP BY DIMITRY SCHIDLOVSKY

however, meant that any intervening layer subsequently discovered had to incorporate a "prime" symbol to distinguish it. Although other layers were eventually renamed, the D″ nomenclature has endured.

Investigators proposed numerous interpretations to account for the seismic properties of the D″ layer. Unfortunately, there were far too many possible explanations and far too little information to permit a definitive characterization of the layer. Better descriptions of the D″ layer had to wait until the technological breakthroughs of the 1980s. Then, using arrays of recording instruments deployed around the world, seismologists could for the first time collect and process enough data to derive three-dimensional images of Earth's interior. They used seismometers that primarily operate in the range between about one and 0.0003 hertz, or cycles per second. (These acoustic frequencies are far below the range of human hearing, which extends from about 20 to 20,000 hertz.) Seismic tomography is often compared to computed tomographic scans used in medicine. But because it relies on sound waves, seismic tomography is more akin to the ultrasonic imaging done during pregnancy. The main drawback is its resolution: images of features smaller than 2,000 kilometers tend to be smeared out.

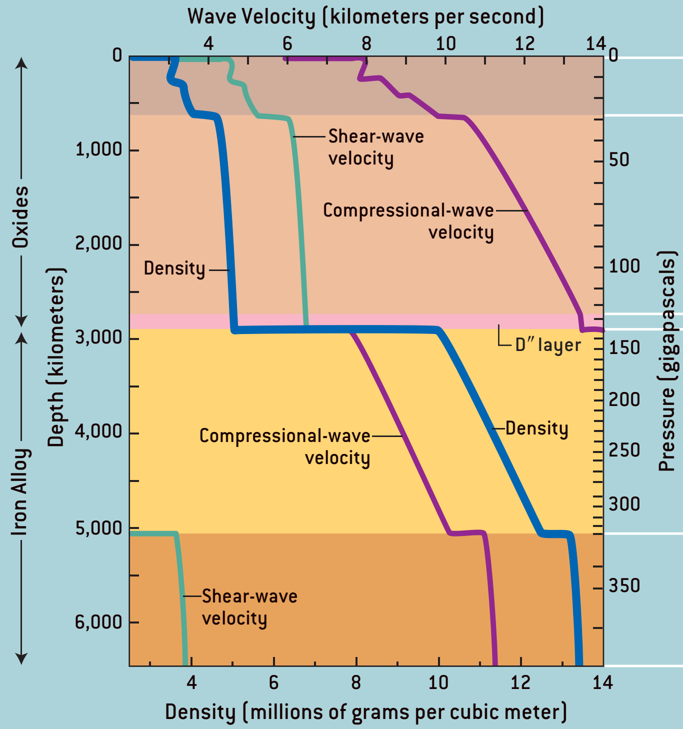
Nevertheless, seismic tomography helped to quantify the properties of the D″ layer. It showed that the region differs drastically from the overlying mantle. The fact that the velocity of seismic waves is affected over continent-size areas shows that large-scale structures dominate D″. Still, seismic tomography could not explain the causes of this variability in physical properties. Could large, chemically distinct structures exist at the bottom of the mantle, just as continents mark the seismic heterogeneity of Earth's surface? Or are the heterogeneities simply large-scale temperature differences at the base of the mantle?

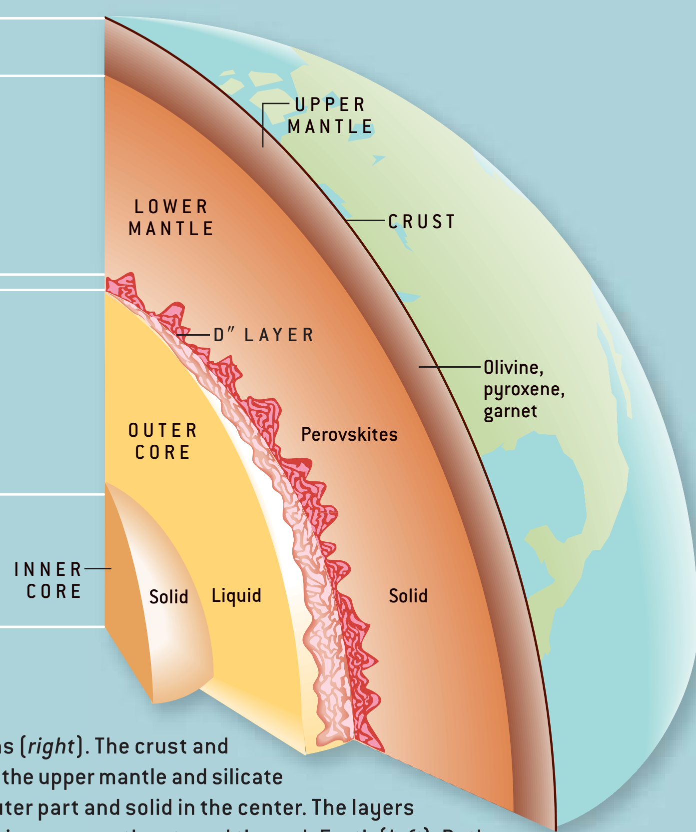
READING THE WAVES

TO ANSWER these questions, one of us (Lay) began in the early 1980s to implement a new method to explore the core-mantle boundary. The idea was to use computer calculations to analyze all the characteristics of the observed seismic wave front, not just the wave velocity, as in the case of seismic tomography. Such waveform analysis is a powerful approach because the technique can resolve structures as small as a few tens of kilometers across instead of those 2,000 kilometers or more in size. The disadvantage is that one can look only at limited parts of the core-mantle boundary. There are not enough earthquakes or other sources of seismic energy to obtain a global picture at such a high level of detail.

The waveform studies suggest that neighboring regions within the D″ layer can be more distinct than had once been thought. For example, several research groups studying the core-mantle

A SEISMIC VOYAGE TO THE CENTER OF EARTH





CROSS SECTION OF EARTH shows the planet's primary regions (*right*). The crust and mantle consist of oxide crystals such as olivine, pyroxene and garnet in the upper mantle and silicate perovskite in the lower mantle. The core is an iron alloy, liquid in the outer part and solid in the center. The layers correspond to the observed variations in density and velocity of seismic waves as they travel through Earth (*left*). Both density and wave velocity increase as a function of depth except at the D″ layer. Note that seismic energy can propagate as shear waves (waves that oscillate at right angles to the direction of motion) and as compressional waves (waves that move back and forth in the travel direction). Because liquids do not have rigidity, shear waves cannot propagate in the outer core. Shear-wave motions reappear in the inner core because a small fraction of the compressional waves transforms into shear waves at the liquid-solid interface.

IAN WORPOLE

boundary below northern Siberia found that acoustic velocities vary so radically over short distances that closely spaced seismometers systematically recorded different waveforms. The finding can best be explained by assuming that the heterogeneity in seismic velocities is large in magnitude and occurs over distances smaller than can be resolved, that is, within a few tens of kilometers. Waveform studies can also map the differences in thickness of the D″ layer. In many places the top of the D″ layer causes an abrupt increase in wave velocity, a process that reflects seismic energy. The reflections have revealed that the thickness of the D″ layer varies dramatically. The layer can be so thin as to be undetectable, or it can span as many as 300 kilometers.

Shear velocities of seismic waves with horizontal and vertical directions of vibration differ within the D″ layer—which is not the case in the overlying lower mantle. That fact suggests a change in rock texture between the D″ region and the overlying mantle. Waveform studies have additionally played an important role in revealing large-scale low-velocity regions in the D″ layer beneath the central Pacific and southern Africa. Lianxing Wen of Stony Brook University and Sidao Ni and Donald Helmberger of the California Institute of Technology have mapped out massive provinces at the base of the mantle composed of apparently chemically distinct material with low shear velocity but normal compressional velocity. These regions underlie low shear velocity regions in the central mantle, and geophysicists have postulated that thermal plumes may rise from the margins of these regions.

Stanley M. Flatté's group at the University of California at Santa Cruz helped to confirm the great variability of the D″ layer. During the mid- to late 1980s, Flatté and his colleagues began to apply new methods of wave analysis to the signals obtained from seismic waves that have been scattered in the deep mantle. Their method relies on a statistical description of how waves propagate through a strongly scattering substance. Such material would be analogous to fog or clouds. Flatté's approach is to observe how the wave front from an earthquake changes shape after traveling through the D″ region. An earthquake initially sends out a smooth, spherically expanding wave. But as that wave is refracted and scattered by variations in seismic features, such as the strong heterogeneities near the core-mantle boundary, the front no longer remains smooth. It becomes rippled, or corrugated [*see illustration on page 41*].

The trick in measuring the degree of wave-front corrugation is a dense array of seismometers. Taking observations from one such collection located in Norway, Flatté has shown that the D″ region appears quite murky to seismic waves. It must contain heterogeneous features as small as 10 kilometers in length. The seismological observations thus indicate that the D″ region is a heterogeneous layer that laterally varies in thickness.

In contrast to the murkiness of the D″ layer, the core-mantle boundary (on which the D″ layer rests) appears smooth and sharp. John E. Vidale, now at the University of California at Los Angeles, and Harley Benz of the U.S. Geological Survey have beautifully demonstrated the abruptness of the interface. They used a vast number of seismic recording stations that had been deployed across the western U.S. The array of seismometers generally monitors regional earthquake activity, but Vidale and Benz have employed it to find seismic waves that have bounced off the core-mantle boundary. Remarkably, seismic waves arrived coherently across more than 900 stations in the array. This coherence implies that the core-mantle boundary represents a sharp transition from the mantle to the core, at least for the area measured. The sudden transition reflects as much as 50 percent of the seismic waves and transmits the remainder. Analyses of the reflected and transmitted waves show that the boundary varies in depth by no more than a few kilometers.

The core-mantle boundary appears to be a sharp reflector, less than a few kilometers in thickness. In some regions, however, a thin ultralow-velocity layer (shear velocities reduced by 15 percent or perhaps more) has been observed using waves reflected from or diffracted along the core-mantle boundary. Helmberger and Edward Garnero of Arizona State University first detected these ultralow-velocity patches, which have the properties expected for partially molten regions adjacent to the core-mantle boundary.

Seismic-wave studies have done much to elucidate the D″ layer and the core-mantle boundary. But the inaccessibility of the regions has prevented geophysicists from understanding completely how such complicated structures came about.

HEAT AND PRESSURE

IF SEISMIC STUDIES cannot thoroughly breach the remoteness of the deep Earth, why not bring the core and mantle to the surface? That is precisely the approach taken by many researchers, including one of us (Jeanloz). Specifically, we sought to duplicate the high pressure and temperature existing in the deep mantle and core. A breakthrough in engineering made such a feat possible: investigators had learned to compress minuscule samples between the points of two diamonds and to heat the specimen using a high-powered laser beam. By 1986 the diamond cells could generate pressures greater than those at Earth's center.

Diamond's hardness is not the only reason for using the substance as an anvil. The utility of diamond also lies in its transparency. A laser beam can be focused directly through the diamond to heat the sample to thousands of degrees Celsius. Moreover, one can observe the specimen while it is at superhigh pressures and temperatures. One determines the temperature of the sample by measuring the thermal radiation the sample emits through the diamond. In this way, one can quantify how "red hot" or "white hot" the material has become; astronomers infer the surface temperatures of stars by color in the same manner. Using the laser-heated diamond cell, we can simulate the appropriate temperatures and pressures at the core-mantle boundary. We wanted to see what would happen when we placed matter that constitutes the outer core in contact with minerals of the lowermost mantle.

Of course, we needed to know what materials make up the mantle and core before squeezing them together. To determine the mantle constituents, Elise Knittle, working with Jeanloz, followed up on research by groups at the Australian National University, the Carnegie Institution of Washington and elsewhere. We relied on prior experimental work, theoretical models and the fact that the pressure in the lower mantle exceeds 20 gigapascals (200,000 atmospheres).

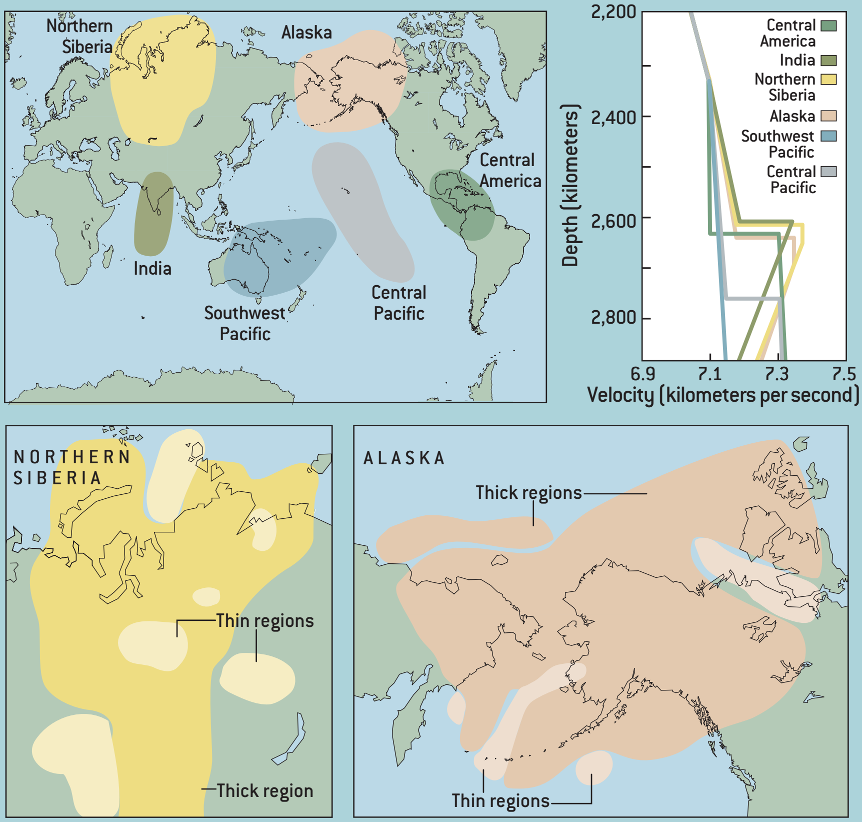
From that information, we deduced that a single high-pressure mineral phase must dominate the lowermost mantle. This mineral is a dense form of iron magnesium silicate, or (Mg,Fe) SiO_3, a robust and chemically simple compound that can be formed only under pressures above 20 gigapascals. Because it has the same crystalline structure as the mineral perovskite ($CaTiO_3$), it is consequently called magnesium silicate perovskite. The lower mantle rock probably also contains minor amounts of magnesiowüstite—a combination of magnesium oxide (MgO) and wüstite (FeO). This composition is quite unlike the nature of rocks at or near Earth's surface. Such surface rocks are composed of many different, complex minerals that react chemically and transform into new minerals under modest

THE AUTHORS

RAYMOND JEANLOZ and *THORNE LAY* study the physics of the deep Earth. Jeanloz, professor of earth and planetary science and of astronomy at the University of California, Berkeley, received his Ph.D. in 1979 from the California Institute of Technology. A MacArthur Fellow, Jeanloz also studies the internal evolution of other terrestrial planets and the formation of new types of glass that have novel properties. Lay is professor of earth sciences at the University of California, Santa Cruz, where he is also director of the Institute of Geophysics and Planetary Physics. His specialty is the study of earthquakes and the structure of Earth's interior. A recipient of the American Geophysical Union's 1991 Macelwane Medal, Lay earned his Ph.D. in 1983 from Caltech.

D IS FOR "DIVERSE"

SHEAR-WAVE VELOCITY in the D″ layer changes across Earth, as indicated by the six regions (*colored areas, top left*) that have been most intensely studied. The corresponding velocity distribution as a function of depth (*top right*) shows that each region exhibits a discontinuity at the D″ layer. The uniqueness of each velocity signature implies that D″ varies over the entire globe. The expanded maps (*bottom*) for areas below northern Siberia and Alaska summarize the heterogeneity of the D″ layer, showing the intermingling of thick regions (*dark patches*) with parts so thin as to be seismically invisible (*light patches*).



IAN WORPOLE

changes of pressure or temperature. The deduced chemical simplicity of the deep mantle accords well with the data derived from seismic waves, which show it to be relatively devoid of structure (except for the D″ layer). This consistency gives us confidence that we are examining the appropriate minerals in our laboratory simulations.

Recent experiments and quantum-mechanical calculations by several groups in Japan and the U.S. indicate that magnesium silicate perovskite may transform to a slightly denser mineral structure at the high pressures and temperatures near the base of the mantle. This transition is expected to occur only in the lower-temperature regions of D″, if it occurs in the D″ layer at all.

Determining the constituent of the core was more straightforward. Seismological studies done more than 50 years ago enabled geophysicists to infer its structure. The core consists of a molten substance surrounding a solid center. The fluid is acknowledged to be a metal—specifically, an alloy of iron. In fact, the churning of the molten iron generates Earth's magnetic field.

Having established the compounds involved, Knittle, now at the University of California at Santa Cruz, carried out a series of experiments in which liquid iron was put in contact with crystalline silicate perovskite at high pressures. She found that the perovskite reacts vigorously with liquid iron, even if these substances touch for just a few seconds. The nature of the chemical reaction is quite interesting and unexpected. The products are a mixture of electrically insulating oxide minerals—magnesium silicate perovskite and stishovite (SiO_2)—and metallic alloys—iron silicide (FeSi) plus wüstite. Wüstite had not been known to be able to form a metallic alloy at any temperature or pressure. Qualitatively speaking, wüstite can react this way because its oxygen atom at high pressures takes on the chemical attributes normally ascribed to its neighbor in the periodic table, sulfur. Metallic sulfides such as iron disulfide (pyrite, or fool's gold) are of course well known.

The experiments also showed that liquid iron begins to react with mantle substances at pressures of 20 to 30 gigapascals. Such pressures are far less than those at the core-mantle boundary (136 gigapascals). Therefore, the reactions have probably persisted since the earliest history of the planet—that is, when Earth was developing and the core might have been forming at pressures below 136 gigapascals. Such chemical reactions are likely to have significantly altered the core-mantle system. A considerable amount of oxygen has probably been drawn into, or alloyed with, the core metal over geologic history. In essence, the lower mantle rock has been and still is slowly dissolving into the liquid metal of the outer core. Berni J. Alder of Lawrence Livermore National Laboratory made this suggestion more than 25 years ago. Our experiments substantiate his conjecture.

Indeed, one of the remarkable consequences of this hypothesis is that it offers a simple explanation for why the properties of the core are nearly but not exactly those of iron at the equivalent pressure and temperature. Most notably, the density of the outer core is about 10 percent lower than that of pure iron [see "The Earth's Core," by Raymond Jeanloz; SCIENTIFIC AMERICAN, September 1983]. But as indicated by Alder's hypothesis and our diamond-cell experiments, the core cannot be completely iron. A purely iron core would have become tainted by reaction with the overlying rock over geologic time. Quite plausibly, the core was never pure iron. Instead it probably contained some nickel, sulfur and other minor constituents. Iron-rich meteorites provide the basis for this hypothesis. Such meteorites, considered partial remnants of the materials from which Earth formed, harbor many similar contaminants. Like pure iron, these iron-rich alloys can react chemically with rocky compounds at high pressures and temperatures, forming an alloy with oxygen.

CORE MEETS MANTLE

ACCORDING TO our experiments, the dense liquid of the outer core must seep into the rock, probably by capillary action. The molten metal would penetrate along the boundaries between the mineral grains at the bottom of the mantle. Estimates of the capillary forces involved suggest that the core liquid could move upward some tens to hundreds of meters above the core-mantle boundary. The reaction between core liquid and mantle rock probably takes place in less than a million years—instantaneously, in geologic terms.

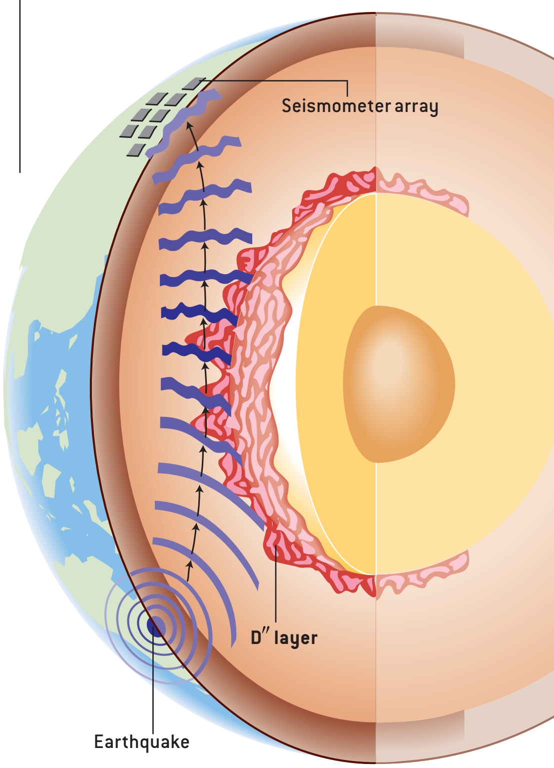
The liquid, however, does not necessarily always have to move upward and to work against gravity. The interface between the mantle and core is not likely to be perfectly flat. Metallic liquid would permeate laterally and downward into the mantle rock from regions where the core-mantle boundary is elevated. Measurements from geodetic and seismological studies indicate that the topography of the core-mantle boundary deviates from absolute flatness by hundreds of meters to a few kilometers. Therefore, the zone of permeation and direct chemical reaction between the core liquid and mantle rock is no more than hundreds to at most thousands of meters thick. The size estimate explains why any reaction zone at the core-mantle boundary is hard to detect. The thickness of the reaction zone is less than typical seismic wavelengths. In addition, no more than a modest fraction of the reaction zone consists of liquid at any given moment. Thus, detecting the presence of a small amount of liquid in a thin region requires very detailed analysis of waves that sense the structure in the lowermost mantle.

How do these chemical reactions at the core-mantle boundary account for the observed characteristics of the D″ layer? The answer lies in a complex and indirect process resulting from forces that act on the core-mantle interface. The forces come from the thermal energy of the underlying core, which heats the rock at the base of the mantle. As a result, the heated part of the mantle moves upward over a period of tens to hundreds of millions of years—far longer than the reaction between the core and mantle, which takes place in less than one million years. The convection must disrupt the reaction zone at the core-mantle boundary, entraining it upward and exposing fresh mantle rock to the corrosive liquid of the core. The convection is the same force that causes the tectonic plates to move at Earth's surface.

Mantle convection does not entrain liquids very far; any liquid metal that might be present in the boundary probably flows out, spongelike, through porous rock before moving upward. On the other hand, the iron-rich crystalline products from the reaction zone, such as wüstite, are readily incorporated into the mantle flow. The slow convection of the mantle pulls up the crystalline alloy a modest distance before the density of the metallic solids causes them to sink back toward the bottom. These solids essentially resemble the dregs of spice that remain at the bottom of a pot of mulled wine.

As a result, the alloy-rich substances would tend to pile up on the bottom of the mantle, especially near regions of upwelling, much as snowdrifts form in a blizzard. The upward dispersal abets infiltration of material from the core and builds a thicker zone of intermixing; the intermixing of reaction products and unreacted mantle causes the seismic heterogeneity. In contrast, downwelling regions would disperse the dregs and thus tend to thin the D″ layer and depress the core-mantle boundary. Modeling by Louise Kellogg of the University of California at Davis and Norman H. Sleep of Stanford University and others suggests that the metallic alloys in local regions of the reaction zone may be swept upward several hundred kilometers into the mantle. The process would require some tens of millions of years.

The buildup of the alloy-rich drifts at the bottom of the mantle solves an important mystery. Specifically, the drifts would explain the variation in thickness of the D″ layer observed by seismologists. Moreover, calculations indicate that the height of the alloy drift swept up in the mantle is comparable to the thickest parts of D″. Given the billions of years for progressive accumulation of the metallic dregs, it is plausible that much of the complexity and many



DISTORTION OF SEISMIC WAVES enables researchers to analyze the heterogeneous characteristics of the D″ layer. Waves emanating from an earthquake are smooth. When they pass through the D″ region, their wave fronts become rippled, or corrugated. The corrugation is measured by a dense array of seismometers located on another part of Earth. One such array, in Norway, was originally constructed to monitor seismic waves generated by underground nuclear tests.

IAN WORPOLE

of the variations in thickness of D″ result from the way mantle flow modulates the alloy-rich reaction layer. The flow may have also caught in its wake other dense mantle material or products from the core. We suspect that reaction dregs can collect, albeit to a lesser extent, on the inner side of the core-mantle boundary. A thinner version of the D″ layer probably exists there, just inside the liquid outer core.

In view of the intense dynamics taking place 2,900 kilometers below Earth's surface, it should not be surprising that the forces in the core-mantle system might be making their presence felt throughout Earth as a whole. Indeed, workers have found tantalizing evidence that suggests that the core-mantle zone strongly influences two features observable at the surface. They are the wobbling in Earth's rotation, known as nutations, and the geomagnetic field.

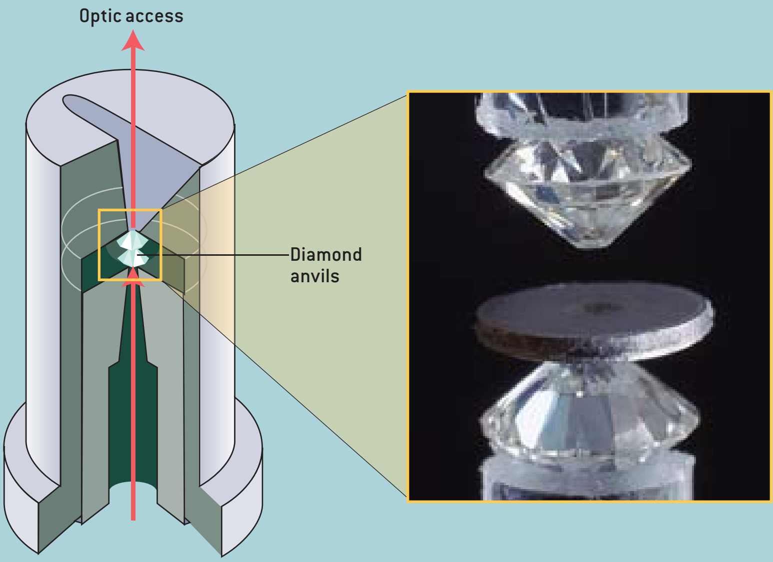
Bruce A. Buffett, now at the University of Chicago, concluded that the core-mantle boundary affects Earth's nutations. He did so after making highly accurate calculations of the wobbling. The workers measured the wobbling using very long baseline interferometry. Radio astronomers often rely on this technique to make highly precise measurements of stellar objects. Various tidal forces had been thought to be solely responsible for Earth's nutations. Such mechanisms include the friction generated as the solid surface of Earth rubs against the atmosphere and oceans as well as the gravitational interactions with the sun and the moon. Buffett discovered, however, a component of the nutations that could not be explained by tidal forces. Motivated by the diamond-cell results, he considered the possibility that a thin reaction zone at the core-mantle boundary might offer an explanation for the anomalous nutation component.

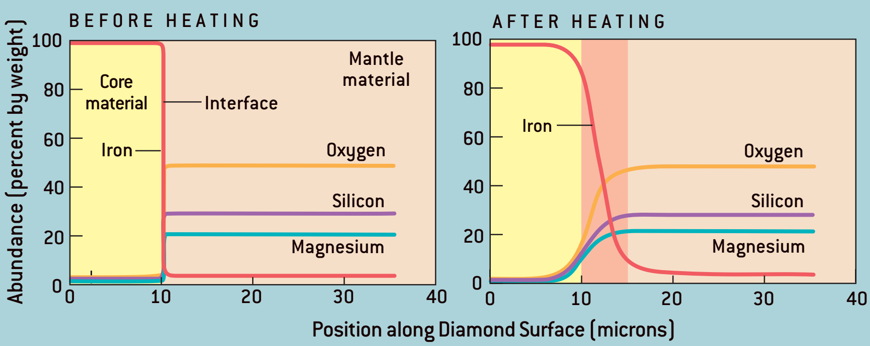
He showed that such a reaction layer can easily account for the nutation signal if the layer contains electrically conducting material, as inferred from experiments. The magnetic field lines emanating from the core would induce small electric currents to flow in the conducting mixture. These currents in turn produce their own magnetic fields. The small magnetic fields interact with the main geomagnetic field lines, much as poles of a magnet can either attract or repel. In essence, the core and mantle behave as two magnets that push against each other. This coupling affects the nutations. The baseline interferometry data are nicely explained if one invokes a heterogeneous reaction zone that contains metal and is a few hundred meters thick.

Indeed, our experiments predicted just such a configuration for the reaction zone. The products of the reaction at the bottom of the mantle are expected to consist of a few tens of percent of electrically conducting alloys, such as iron silicide and wüstite. A zone consisting of only 15 to 20 percent alloy would be sufficient to account for the nutations. Thus, our conclusion that the reaction zone would be hundreds of meters thick and would fluctuate in thickness and conductivity along the

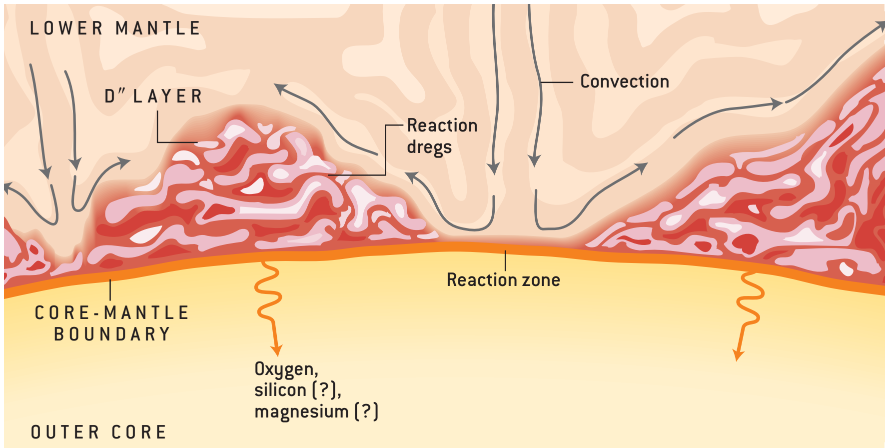
CONJURING CORE CHEMICALS

DIAMOND-ANVIL high-pressure cell (*below*) can duplicate the pressures and temperatures of the deep Earth. The material to be squeezed and heated is placed in a metal-foil gasket between the tips of two diamond anvils (*photograph*). Turning a thumbscrew (*not shown*) brings the anvils together, compressing the sample. A laser beam can be focused through the diamond to heat the sample. Compositional profiles (*bottom*) show the abundance of iron, oxygen, silicon and magnesium (elements at the core-mantle boundary) across a sample before and after heating. When heated, the interface region between iron and mantle material broadens, spanning a width of about five microns. The broadening is caused by chemical reactions that produce a mixture of metallic alloys (FeSi and FeO) and insulating oxides ($MgSiO_3$ and SiO_2).





IAN WORPOLE (*upper left and bottom*); DOUGLAS L. PECK (*photograph*)



D″ LAYER forms as a result of chemical reactions between the core and mantle. In essence, the mantle rock partly dissolves in the liquid iron of the outer core, producing metal-rich "dregs" that are deposited on the core-mantle boundary. Convection in the mantle tends to disperse the products under downwelling regions and to build up material at upwellings. A thin layer enriched in oxygen and possibly silicon and magnesium may exist on the inner side of the core-mantle interface.

IAN WORPOLE

core-mantle boundary accords well with Buffett's hypothesis.

The second observable surface effect that the core-mantle region influences is Earth's magnetic field. The origin of the main geomagnetic field is well understood, at least in general terms. A dynamo effect, rather than conventional magnetism of the iron in the core, produces the geomagnetic field. (Iron is no longer magnetic at either the pressures or temperatures existing in the core.) Churning of the liquid-metal outer core creates electric currents, analogous to a wire moved through an electric field. Like a wire carrying a current, the core then generates a magnetic field around itself.

Convection powers the motion of the molten outer core. The hot liquid from deep inside rises toward the cooler top of the core. The movement transfers heat upward and causes a convective flow. Cooler liquid from near the core-mantle boundary sinks downward and thus also helps to power the convection. Additional sources of convection, such as internal separation of solids and liquid in the outer core, are possible. In this way, the mechanical energy of convection—fluid flow in the outer core—is converted to magnetic energy.

REVERSING FIELD

THE PRINCIPLES that govern this process are called magnetohydrodynamics—a combination of hydrodynamics, or the physics of fluid flow, and electromagnetism. The mathematical equations behind the process, however, are so complicated that no one has been able to solve them in complete generality. As a result, the solutions obtained are based on physically plausible but greatly simplified assumptions. The solutions obtained from these assumptions do not necessarily explain the small but observable details of Earth's magnetic field, such as the slight ripples in the field intensity. Perhaps the discrepancy results from one of the traditional simplifications used in the calculation: that the metallic core is surrounded by an electrically insulating region, corresponding to the mantle. Geophysicists are now recognizing that the lowermost mantle is not completely insulating but consists of a mixture of metallic alloys and insulating silicates.

Motivated by this information, Friedrich H. Busse of Bayreuth University in Germany reexamined the magnetohydrodynamic equations. He discovered a new class of mathematical solutions to the dynamo problem that result directly from the variations in electrical conductivity in the lowermost mantle. The solutions depend on two major factors. One is that the geomagnetic field lines are essentially "frozen" into the liquid metal of the outer core. So, locked into place, the field lines move only with the convective flow of the liquid outer core. The second factor is that metallic regions embedded within the D″ layer interfere with the horizontal movement of magnetic field lines emanating from the core. The D″ layer can then deflect or pile together the field lines from the core. Both factors would, according to Busse's calculations, create local magnetic fields at the bottom of the mantle. The fields would explain several complexities of the geomagnetic field, including the observed ripples in field strength.

The electromagnetic characteristics of the core-mantle boundary may affect the reversals of Earth's magnetic field [see "Probing the Geodynamo," on page 28]. During reversals, which occur every few 100,000 years, the magnetic poles seem to follow a preferred trajectory. Such preference is especially evident for the most recent reversals in Earth's history. The late S. Keith Runcorn of Imperial College London and the University of Alaska–Fairbanks postulated several mechanisms by which the electrical variations of the D″ layer might influence the path of the magnetic poles.

In a sense, then, the dynamics between the core and mantle extend beyond Earth, stretching well into space via the geomagnetic field. We now recognize the planetary importance of the core-mantle interface, and improved technology is certain to clarify how this remote region shapes the evolution of Earth. SA

MORE TO EXPLORE

The Nature of the Earth's Core. R. Jeanloz in *Annual Review of Earth and Planetary Sciences,* Vol. 18, pages 357–386; 1990.

Earth's Core-Mantle Boundary: Results of Experiments at High Pressures and Temperatures. E. Knittle and R. Jeanloz in *Science,* Vol. 251, pages 1438–1443; March 22, 1991.

Deep Interior of the Earth. John A. Jacobs. Chapman & Hall, 1992.

The Core-Mantle Boundary Layer and Deep Earth Dynamics. T. Lay, Q. Williams and E. J. Garnero in *Nature,* Vol. 392, pages 461–468; April 1998.

The Evolution of Continental Crust

The high-standing continents owe their existence to Earth's long history of plate-tectonic activity

MOON

Except perhaps for some remote island dwellers, most people have a natural tendency to view continents as fundamental, permanent and even characteristic features of Earth. One easily forgets that the world's continental platforms amount only to scattered and isolated masses on a planet that is largely covered by water. But when viewed from space, the correct picture of Earth becomes immediately clear. It is a blue planet. From this perspective it seems quite extraordinary that over its long history Earth could manage to hold a small fraction of its surface always above the sea—enabling, among other things, human evolution to proceed on dry land.

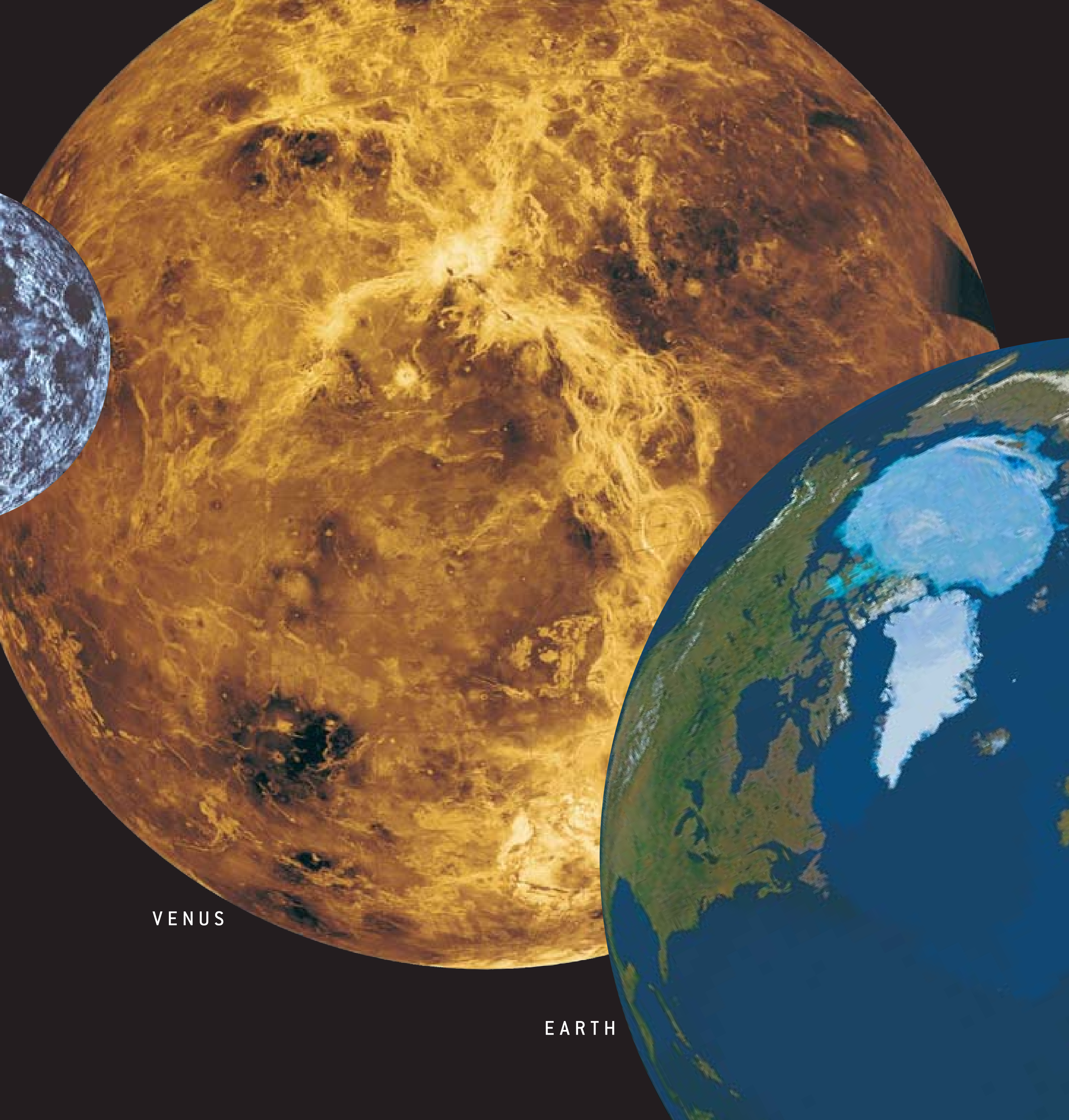
Is the persistence of high-standing continents just fortuitous? How did Earth's complicated crust come into existence? Has it been there all the time, like some primeval icing on a planetary cake, or has it evolved through the ages? Such questions had engendered debates that divided scientists for many decades, but the fascinating story of how the terrestrial surface came to take its present form is now essentially resolved. That understanding shows, remarkably enough, that the conditions required to form the continents of Earth may be unmatched in the rest of the solar system.

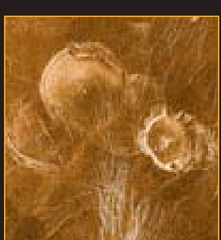
By S. Ross Taylor and Scott M. McLennan

Earth and Venus, being roughly the same size and distance from the sun, are often regarded as twin planets. So it is natural to wonder how the crust of Venus compares with that of our own world. Although centuries of telescopic observations from Earth could give no insight, beginning in 1990 the Magellan space probe's orbiting radar penetrated the thick clouds that enshroud Venus and revealed its surface with stunning clarity. From the detailed images of landforms, planetary scientists can surmise the type of rock that covers Venus.

Our sister planet appears to be blanketed by rock of basaltic composition—much like the dark, fine-grained rocks that line the ocean basins on Earth. Magellan's mapping, however, failed to find extensive areas analogous to Earth's continental crust. Elevated regions named Aphrodite Terra and Ishtar Terra appear to be remnants of crumpled basaltic lavas. Smaller, dome-shaped mounds are found on Venus, and these forms might indicate that volcanic rocks with the composition of granite do exist in some places, but radar reflections show that these pancake-like features may be composed merely of more basalt.

After analyzing the wealth of radar data provided by Magellan, scientists have concluded that plate tectonics (that is, the continual creation, motion and destruction of parts of the planet's surface) does not seem to operate on Venus. There are no obvious equivalents to the extensive mid-ocean ridges or to the great trench systems of Earth.





EARTH'S CRUST is composed primarily of the basaltic rocks that line the ocean basins. Granitic rocks constitute the high-standing continental platforms. Venus is nearly the same size as Earth, yet radar imagery indicates that it is encrusted almost entirely by basalt. Only a tiny fraction of that planet's surface exhibits pancake-shaped plateaus (*detail at left*) that might, like Earth's continents, be built of granitic material. The crust of Earth's moon is largely covered by white highlands formed as that body first cooled from a molten state; volcanic eruptions later created dark so-called seas of basalt.

Thus, it is unlikely that the crust of Venus regularly recycles back into that planet's mantle. Nor would there seem to be much need to make room for new crust: the amount of lava currently erupting on Venus is roughly equivalent to the output of one Hawaiian volcano, Kilauea—a mere dribble for the planet as a whole. These findings from Venus and similar surveys of other solid bodies in the solar system show that planetary crusts can be conveniently divided into three fundamental types.

So-called primary crusts date back to the beginnings of the solar system. They emerged after large chunks of primordial material came crashing into a growing planet, releasing enough energy to cause the original protoplanet to melt. As the molten rock began to cool, crystals of some types of minerals solidified relatively early and could separate from the body of magma. This process, for example, probably created the white highlands of the moon after low-density grains of the mineral feldspar floated to the top of an early lunar "ocean" of molten basalt. The crusts of many satellites of the giant outer planets, composed of mixtures of rock with water, methane and ammonia ices, may also have arisen from catastrophic melting during initial accretion.

In contrast to the product of such sudden, large-scale episodes of melting, secondary crusts form after heat from the decay of radioactive elements gradually accumulates within a planetary body. Such slow heating causes a small fraction of the planet's rocky mantle to melt and usually results in the eruption of basaltic lavas. The surfaces of Mars and Venus and Earth's ocean floors are covered by secondary crusts created in this way. The lunar maria (the "seas" of the ancient astronomers) also formed from basaltic lavas that originated deep in the moon's interior. Heat from radioactivity—or perhaps from the flexing induced by tidal forces—on some icy moons of the outer solar system may, too, have generated secondary crusts.

Unlike these comparatively common types, so-called tertiary crust may form if surface layers are returned back into the mantle of a geologically active planet. Like a form of continuous distillation, volcanism can then lead to the production of highly differentiated magma of a composition that is distinct from basalt—closer to that of the light-colored igneous rock granite. Because the recycling necessary to generate granitic magmas can occur only on a planet where plate tectonics operates, such a composition is rare in the solar system. The formation of continental crust on Earth may be its sole location.

Despite the small number of examples within each category, one generalization about the genesis of planetary surfaces seems easy to make: there are clear differences in the rates at which primary, secondary and tertiary crusts form. The moon, for instance, generated its white, feldspar-rich primary crust—about 9 percent of lunar volume—in only a few million years. Secondary crusts evolve much more slowly. The moon's basalt maria (secondary crust) are just a few hundred meters thick and make up a mere one tenth of 1 percent of the moon's volume, and yet these so-called seas required more than a billion years to form. Another example of secondary crust, the basaltic oceanic basins of our planet (which constitute about one tenth of 1 percent of Earth's mass), formed over a period of about 200 million years. Slow as these rates are, the creation of tertiary crust is even less efficient. Earth has taken several billion years to produce its tertiary crust—the continents. These features amount to just about one half of 1 percent of the mass of the planet.

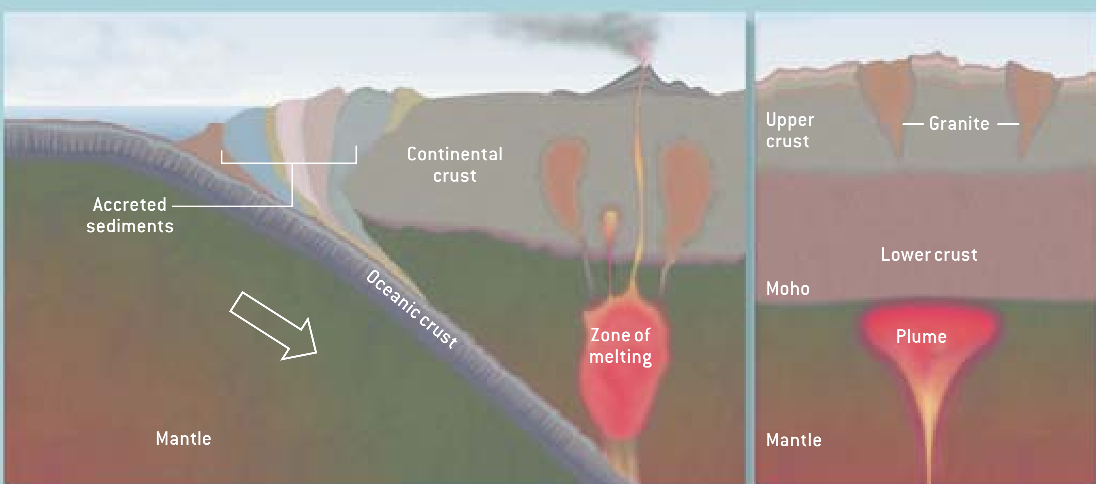
FLOATING CONTINENTS

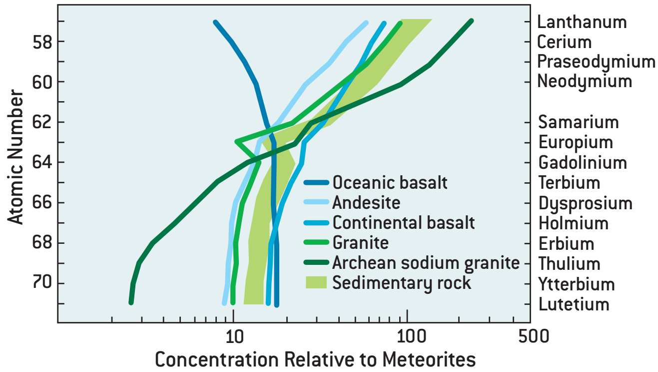
MANY ELEMENTS that are otherwise rarely found on Earth are enriched in granitic rocks, and this phenomenon gives the continental crust an importance out of proportion to its tiny mass. But geologists have not been able to estimate the overall composition of crust—a necessary starting point for any investigation of its origin and evolution—by direct observation. One conceivable method might be to compile existing descriptions of rocks that outcrop at the surface. Even this

PRECEDING PAGES: JET PROPULSION LABORATORY/NASA (*detail*); JPL/NASA (*Venus*); TOM VAN SANT *Geosphere Project* (*Earth*); NASA (*moon*); THIS PAGE: MICHAEL GOODMAN

TECTONICS TOASTS UP THE CRUST

PLATE-TECTONIC ACTIVITY carries oceanic crust deep into Earth (*right*), burying wet sediments along with the descending slab. At a depth of 80 kilometers, high temperatures drive water from the sediments, inducing the overlying rock to melt. The magma generated in this process rises buoyantly and forms new continental material near the surface. As this crust matures (*far right*), heat from radioactivity (or from rising plumes of basaltic magma) may trigger melting at shallow levels to produce granitic intrusions. Such episodes create an upper layer of continental crust that is made up largely of granite.





RARE-EARTH ELEMENT abundance patterns provide characteristic chemical markers for the types of rock that have formed Earth's crust. Although igneous rocks (those that solidify from magma) can have highly variable rare-earth element signatures (*thin lines*), the pattern for most sedimentary rocks falls within a narrow range (*thick band*). That uniformity arises because sediments effectively record the average composition of the upper continental crust.

large body of information might well prove insufficient. A large-scale exploration program that could reach deeply enough into the crust for a meaningful sample would press the limits of modern drilling technology and would, in any event, be prohibitively expensive.

Fortunately, a simpler solution is at hand. Nature has already accomplished a widespread sampling through the erosion and deposition of sediments. Lowly muds, now turned into solid sedimentary rock, give a surprisingly good average composition for the exposed continental crust. These samples are, however, missing those elements that are soluble in water, such as sodium and calcium. Among the insoluble materials that are transferred from the crust into sediments without distortion in their relative abundances are the 14 rare-earth elements, known to geochemists as REEs. These elemental tags are uniquely useful in deciphering crustal composition because their atoms do not fit neatly into the crystal structure of most common minerals. They tend instead to be concentrated in the late-forming granitic products of a cooling magma that make up most of the continental crust.

Because the REE patterns found in a variety of sediments are so similar, geochemists surmise that weathering, erosion and sedimentation must mix different igneous source rocks efficiently enough to create an overall sample of the continental crust. All the members of the REE group establish a signature of upper crustal composition and preserve, in the shapes of the elemental abundance patterns, a record of the igneous events that may have influenced the makeup of the crust.

Using these geochemical tracers, geologists have, for example, determined that the composition of the upper part of the continental crust approximates that of granodiorite, an ordinary igneous rock that consists largely of light-colored quartz and feldspar, along with a peppering of various dark minerals. Deep within the continental crust, below about 10 to 15 kilometers, rock of a more basaltic composition is probably common. The exact nature of this material remains controversial, and geologists are currently testing their ideas using measurements of the heat produced within the crust by the important radioactive elements uranium, thorium and ^{40}K, the radioactive isotope of potassium. But it seems reasonable that at least parts of this inaccessible and enigmatic region may consist of basalt trapped and underplated beneath the lower-density continents.

It is this physical property of granitic rock—low density—that explains why most of the continents are not submerged. Continental crust rises on average 125 meters above sea level, and some 15 percent of the continental area extends over two kilometers in elevation. These great heights contrast markedly with the depths of ocean floors, which average about four kilometers below sea level—a direct consequence of their being lined by dense oceanic crust composed mostly of basalt and a thin veneer of sediment.

At the base of the crust lies the so-called Mohorovicic discontinuity (a tongue-twisting name geologists invariably shorten to "Moho"). This deep surface marks a radical change in composition to an extremely dense rock rich in the mineral olivine that everywhere underlies both oceans and continents. Geophysical studies using seismic waves have traced the Moho worldwide. Such research has also indicated that the mantle below the continents may be permanently attached at the top. These relatively cool subcrustal "keels" can be as much as 400 kilometers thick and appear to ride with the continents during their plate-tectonic wanderings. Support for this notion comes from the analysis of tiny mineral inclusions found within diamonds, which are thought to originate deep in this subcrustal region. Measurements show that diamonds can be up to three billion years old and thus demonstrate the antiquity of the deep continental roots.

It is curious to reflect that less than 50 years ago, there was no evidence that the rocks lining ocean basins differed in any fundamental way from those found on land. The oceans were simply thought to be floored with foundered or sunken continents. This perception grew naturally enough from the concept that the continental crust was a world-encircling feature that had arisen as a kind of scum on an initially molten planet. Although it now appears certain that Earth did in fact melt very early, it seems that a primary granitic crust, of the type presumed decades ago, never actually existed.

THE AUTHORS

S. ROSS TAYLOR and *SCOTT M. McLENNAN* have worked together since 1977 examining Earth's crustal evolution. Taylor has also actively pursued lunar and planetary studies and has published many books on planetology. He is a foreign associate of the National Academy of Sciences. Taylor is currently with the department of earth and marine sciences at the Australian National University and the Lunar and Planetary Institute in Houston. McLennan is professor in the department of geosciences at Stony Brook University. His research applies the geochemistry of sedimentary rocks to studies of crustal evolution on Earth and Mars. McLennan is a member of the Mars Exploration Rover science team.

LUCY READING-IKKANDA

THE EVOLUTION OF GEODIVERSITY

HOW WAS IT that two such distinct kinds of crust, continental and oceanic, managed to arise on Earth? To answer this question, one needs to consider the earliest history of the solar system. In the region of the primordial solar nebula occupied by Earth's orbit, gas was mostly swept away, and only rocky debris large enough to survive intense early solar activity accumulated. These objects themselves must have grown by accretion, before finally falling together to form our planet, a process that required about 50 million to 100 million years.

Late in this stage of formation, a massive planetesimal, perhaps one the size of Mars, crashed into the nearly fully formed Earth. The rocky mantle of the impactor was ejected into orbit and became the moon while the metallic core of the body fell into Earth. As might be expected, this event proved catastrophic: it totally melted the newly formed planet. As Earth later cooled and solidified, an early basaltic crust probably formed.

It is likely that at this stage the surface of Earth resembled the current appearance of Venus; however, none of this primary crust has survived. Whether it sank into the mantle in a manner similar to that taking place on Earth or piled up in localized masses until it was thick enough to transform into a denser rock and sink remains uncertain. In any event, there is no evidence of substantial granitic crust at this early stage. Telltale evidence of such a crust should have survived in the form of scattered grains of the mineral zircon, which forms within granite and is very resistant to erosion. Although a few ancient zircons dating from near this time have been found (the oldest examples are from sedimentary rocks in Australia and are about 4.3 billion years old), these grains are exceedingly scarce.

SEDIMENT pouring from the mouth of the Amazon River into the Atlantic Ocean is derived from erosion within the river's drainage basin and therefore provides an average sampling of large regions of the exposed continental crust of Earth. Analyzing such sediment can yield estimates of the average composition of the upper continental crust, especially the relatively insoluble elements such as the rare-earth elements. Buried sediments turn into sedimentary rocks that are preserved over most of Earth's history, making it possible to trace the evolution of the upper continental crust over billions of years.

NASA/GSFC/JPL, MISR TEAM

More information about the early crust comes from the most ancient rocks to have survived intact. These rocks formed deep within the crust just less than four billion years ago and now outcrop at the surface in northwest Canada. This rock formation is called the Acasta Gneiss. Slightly younger examples of early crust have been documented at several locations throughout the world, although the best studied of these ancient formations is in western Greenland. The abundance of sedimentary rock there attests to the presence of running water and to the existence of what were probably true oceans during this remote epoch. But even these extraordinarily old rocks from Canada and Greenland date from some 400 million to 500 million years after the initial accretion of Earth, a gap in the geologic record caused, no doubt, by massive impacts that severely disrupted Earth's earliest crust.

From the record preserved in sedimentary rocks, geologists know that the formation of continental crust has been an ongoing process throughout Earth's long history. But the creation of crust has not always had the same character. For example, at the boundary between the Archean and Proterozoic eons, around 2.5 billion years ago, a distinct change in the rock record occurs. The composition of the upper crust before this break contained less evolved constituents, composed of a mixture of basalt and sodium-rich granites. These rocks make up the so-called tonalite-trondjemite-granodiorite, or TTG, suite. This composition differs considerably from the present upper crust, which is dominated by potassium-rich granites.

The profound change in crustal composition 2.5 billion years ago appears to be linked to changes in Earth's tectonic regime. Before this time, higher levels of radioactive decay produced more heat in the planet. The consequence was that in the earlier Archean the oceanic crust was hotter, thicker and more buoyant and was not able to be subducted. Instead, under thicker sections of crust that may resemble modern Iceland, denser crust melted and produced the sodium-rich igneous rocks of the TTG suite.

Somewhat similar rocks now form in a few places such as southern Chile, where young oceanic crust subducts. But these modern rocks, forming now because of plate tectonics, are subtly different from their older Archean cousins, which formed from sinking slabs under thick crust. Modern-style plate tectonics did not begin operating until the late Archean (between 3.0 billion and 2.5 billion years ago), when the oceanic crust became cooler, lost its buoyancy and was thus able to sink back into the mantle.

The early tendency for magma to form with a TTG composition explains why crust grew as a mixture of basalt and tonalite during the Archean eon. Large amounts—at least 50 percent and perhaps as much as 70 percent of the continental crust—emerged at this time, with a major episode of growth between 3.0 billion and 2.5 billion years ago. Since that time, the relative height of ocean basins and continental platforms has remained comparatively stable. With the onset of the Proterozoic eon 2.5 billion years ago, the crust had already assumed much of its present makeup, and modern plate-tectonic cycling began.

Currently oceanic crust forms by the eruption of basaltic lava along a globe-encircling network of mid-ocean ridges. More than 18 cubic kilometers of rock are produced every year by this process. The slab of newly formed crust rides on top of an outer layer of the mantle, which together make up the rigid lithosphere. The oceanic lithosphere sinks back into the mantle at so-called subduction zones, which leave conspicuous scars on the ocean floor in the form of deep trenches. At these sites the descending slab of lithosphere carries wet marine sediments as well as basalt plunging into the mantle.

At a depth of about 80 kilometers, heat drives water and other volatile components from the subducted sediments into the overlying mantle. These substances then act as a flux does at a foundry, inducing melting in the surrounding material at reduced temperatures. The magma fractionates, producing andesites, while the more basic substratum probably sinks back into the mantle in a process called delamination. The andesite magma produced in this fashion eventually reaches the surface, where it causes spectacular, explosive eruptions. The 1980 eruption of Mount St. Helens is an example of such a geologic cataclysm. Great chains of volcanoes—such as the Andes—powered by boiling volatiles add on average about two cubic kilometers of lava and ash to the continents every year. This andesite provides the bulk material of the continents.

But the more silica-rich granitic rock, which we see at the surface of the continents, comes from within the crust. The accumulation of heat deep within the continental crust itself can cause melting, and the resultant magma will ultimately migrate to the surface. Although some of this necessary heat might come from the decay of radioactive elements, a more likely source is basaltic magma that rises from deeper in the mantle and becomes trapped under the granitic lid; the molten rock then acts like a burner under a frying pan.

CRUSTAL GROWTH SPURTS

ALTHOUGH THE MOST DRAMATIC SHIFT in the generation of continental crust happened at the end of the Archean eon, 2.5 billion years ago, the continents appear to have experienced episodic changes throughout all of geologic time. For example, sizable, later additions to the continental crust occurred from 2.0 to 1.7, from 1.3 to 1.1 and from 0.5 to 0.3 billion years ago. That Earth's continents experienced such a punctuated evolution might appear at first to be counterintuitive. Why, after all, should crust form in spurts if the generation of internal heat—and its liberation through crustal recycling—is a continuous process?

A more detailed understanding of plate tectonics helps to solve this puzzle. During the Permian period (about 250 million years ago), the major continents of Earth converged to create one enormous landmass called Pangaea [see "Earth before Pangaea," on page 14]. This configuration was not unique. The formation of such "supercontinents" appears to recur at intervals of about 600 million years. Major tectonic cycles driving the continents apart and together have been documented as far back as the Early Proterozoic, and there are even suggestions that the first supercontinent may have formed earlier, during the Archean.

VOLCANOES, such as this one erupting on Russia's Kamchatka Peninsula, mark where new continental material forms above subducting oceanic crust. Over thousands of years, such geologically active terranes will evolve into stable continental crust.

NASA

Such large-scale tectonic cycles serve to modulate the tempo of crustal growth. When a supercontinent breaks itself apart, oceanic crust is at its oldest and hence most likely to form new continental crust after it subducts. As the individual continents reconverge, volcanic arcs (curved chains of volcanoes created near subduction zones) collide with continental platforms. Such episodes preserve new crust as the arc rocks are added to the margins of the continents.

For more than four billion years, the peripatetic continents have assembled themselves in fits and starts from many disparate terranes. Buried in the resulting amalgam is the last remaining testament available for the bulk of Earth's history. That story, assembled from rocks that are like so many jumbled pieces of a puzzle, has taken some time to sort out. But the understanding of crustal origin and evolution is now sufficient to show that of all the planets Earth appears truly exceptional. By a fortunate accident of nature—the ability to maintain plate-tectonic activity—one planet alone has been able to generate the sizable patches of stable continental crust that we find so convenient to live on. SA

MORE TO EXPLORE

No Water, No Granites—No Oceans, No Continents. I. H. Campbell and S. R. Taylor in *Geophysical Research Letters,* Vol. 10, No. 11, pages 1061–1064; November 1983.

The Continental Crust: Its Composition and Evolution. S. Ross Taylor and Scott M. McLennan. Blackwell Scientific Publications, 1985.

The Geochemical Evolution of the Continental Crust. S. Ross Taylor and Scott M. McLennan in *Reviews of Geophysics,* Vol. 33, No. 2, pages 241–265; May 1995.

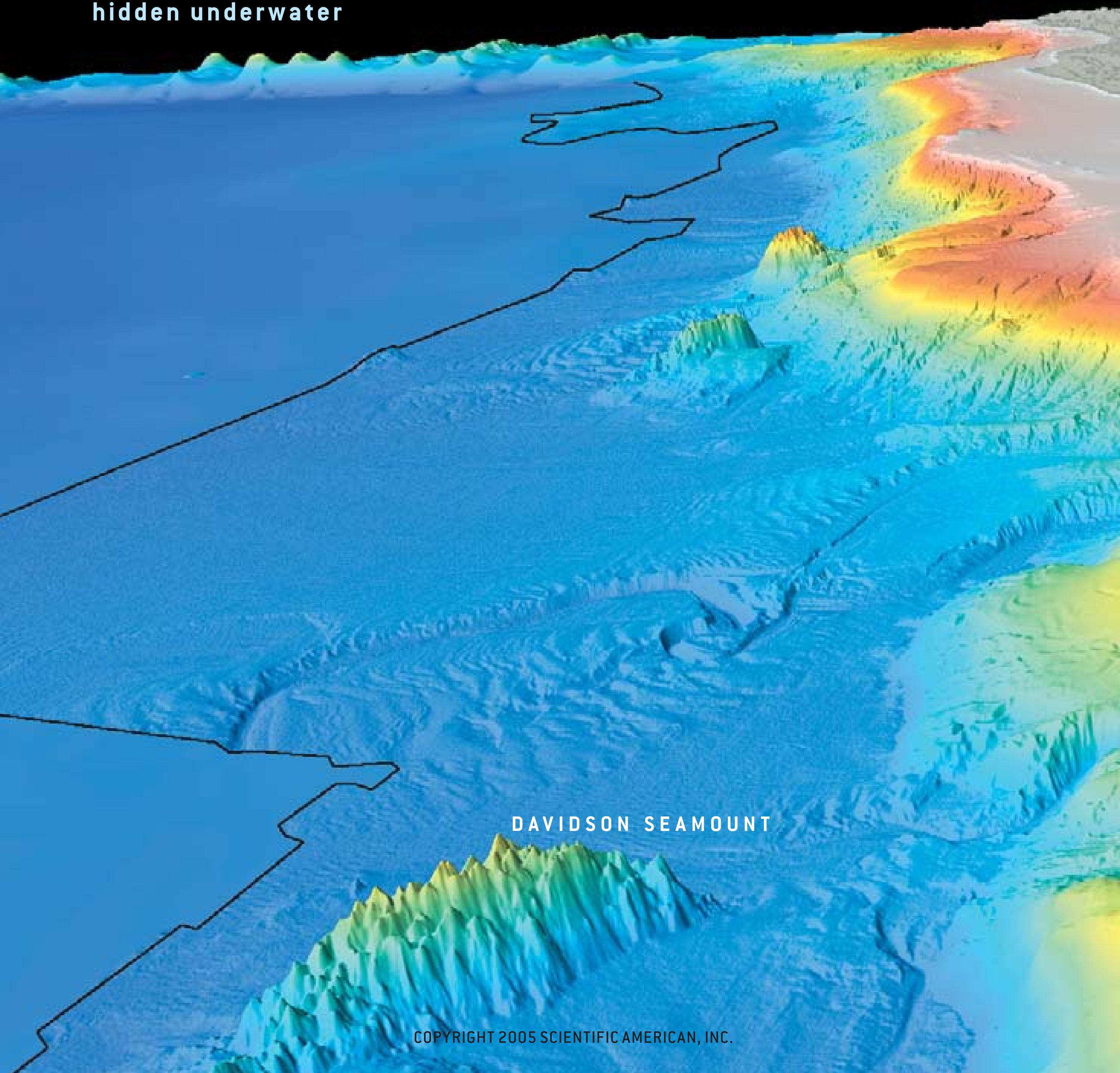
Nature and Composition of the Continental Crust: A Lower Crustal Perspective. Roberta L. Rudnick and David M. Fountain in *Reviews of Geophysics,* Vol. 33, No. 3, pages 267–309; August 1995.

Earth as an Evolving Planetary System. Kent C. Condie. Academic Press, 2004.

Evolution and Differentiation of the Continental Crust. Edited by Michael Brown and Tracy Rushmer. Cambridge University Press, 2005.

Panoramas of the

Modern sonar techniques map the continental margins of the U.S. and reveal the richly varied scenery usually hidden underwater

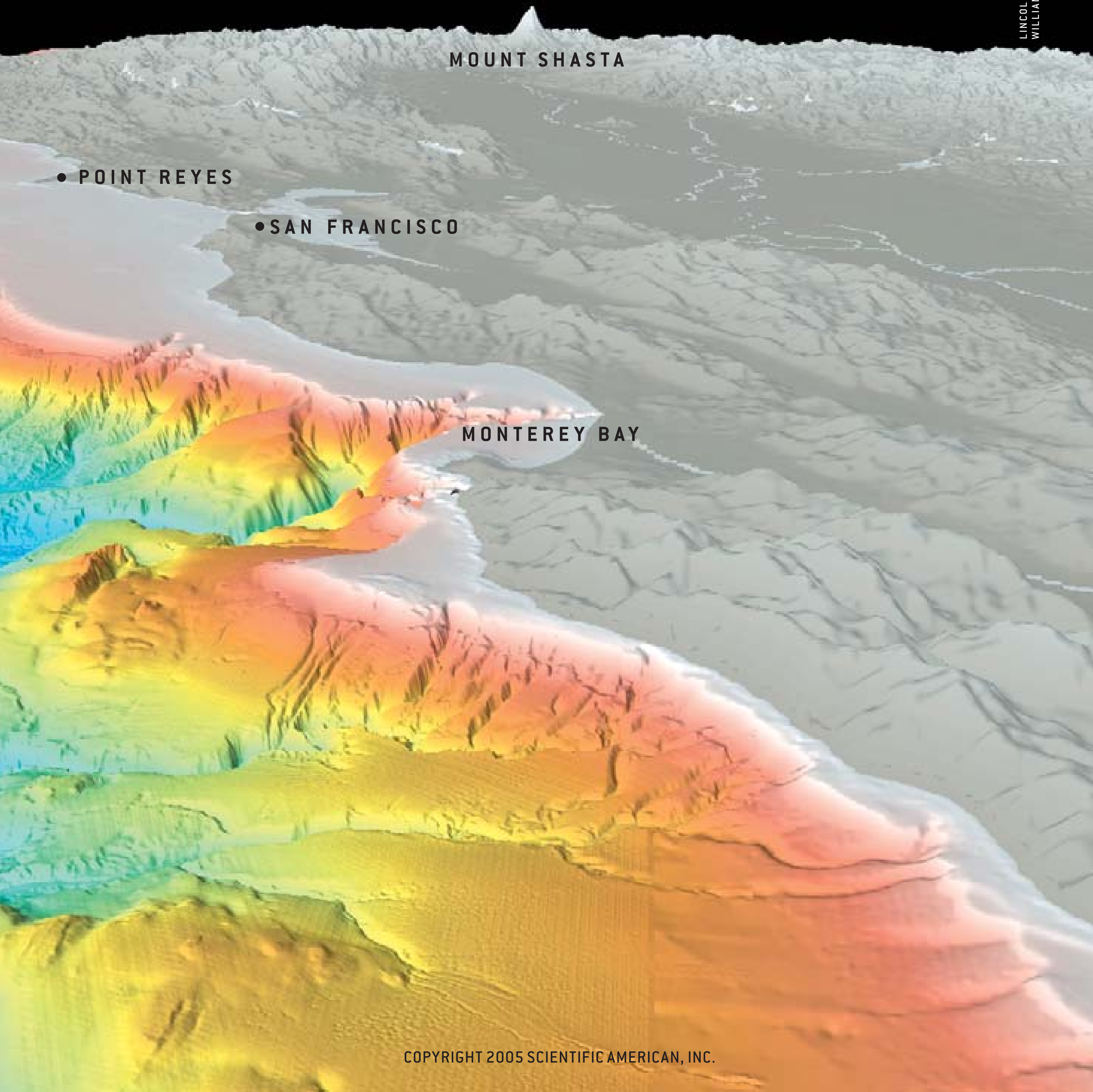


Seafloor

By Lincoln F. Pratson and William F. Haxby

COMPUTER-GENERATED IMAGES of the seafloor surrounding the U.S. show geologic features in great detail in regions where specialized sonar mapping has been done (*right of black line*).

LINCOLN F. PRATSON AND WILLIAM F. HAXBY



In 85 B.C. or thereabouts, a Greek named Posidonius set sail on a curious mission. He was not carrying freight or passengers, nor was he engaged in warfare. He simply wanted to answer an age-old question: How deep is the ocean? Halting his vessel in the middle of the Mediterranean Sea, Posidonius coaxed his ship's crew to let out nearly two kilometers of rope before a large stone attached to the end of the line finally hit bottom. He and his men must have been jubilant—at least until they realized that they then had to haul the great weight back onboard.

For the next 2,000 years, naval surveyors and oceanographers continued to use exactly the same laborious line-and-sinker method to probe the oceans' depths. It is not surprising that they made scant progress. Then, during the 1920s, oceanographers developed the first echo sounders—instruments that could measure the deep water by bouncing sound waves off the bottom. With the wealth of measurements these devices provided, scientists got their first glimmers of the true shape of the ocean basins.

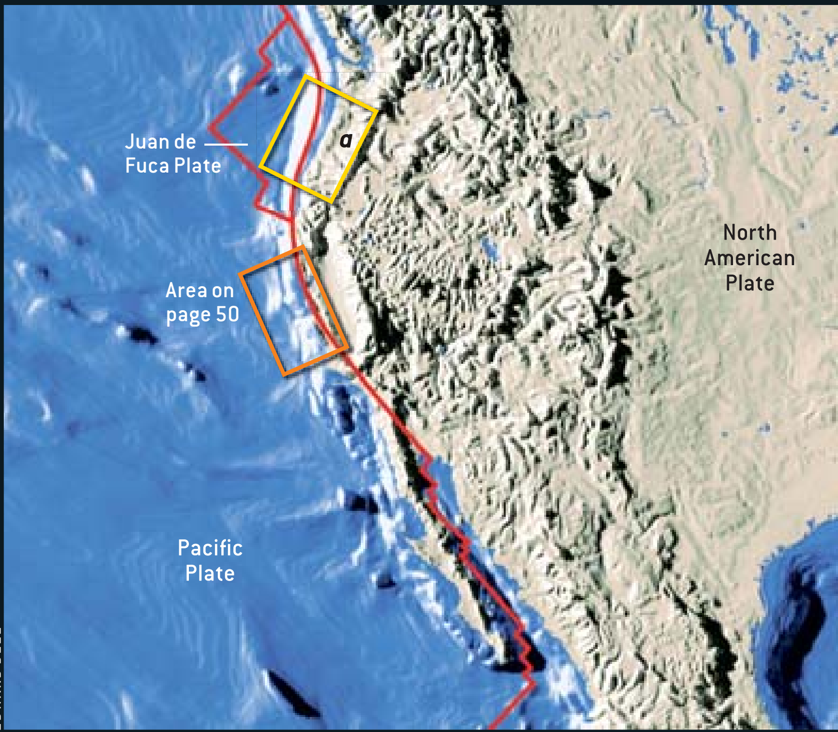
In the past few decades engineers have constructed ever more sophisticated acoustic devices to speed the mapping of this hitherto hidden part of Earth. The major impetus for these developments initially came from concerns about national defense, but more recently economic considerations have taken precedence.

Beginning with the U.S. in 1981, the world's

THE AUTHORS

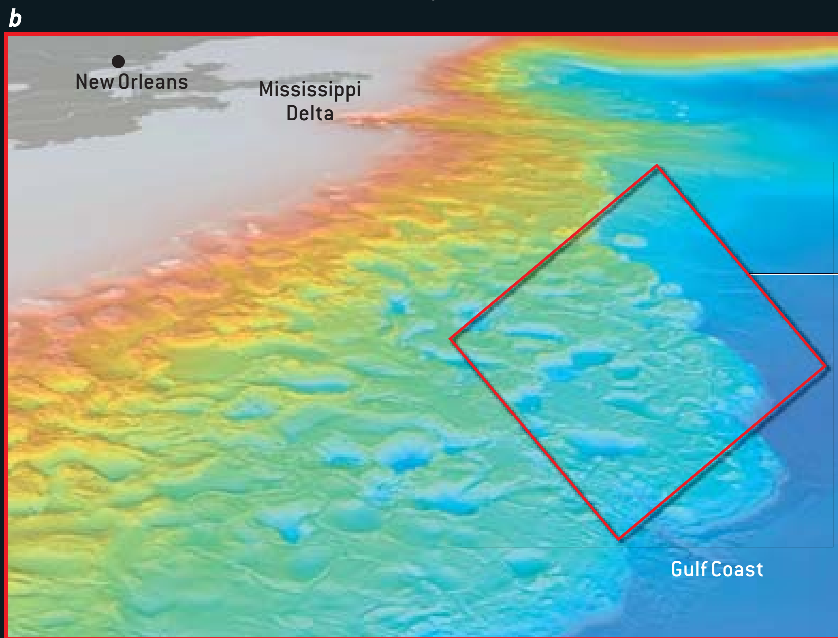
LINCOLN F. PRATSON and *WILLIAM F. HAXBY* have worked together and independently for many years probing the continental margins of the U.S. and elsewhere in the world. Pratson completed his Ph.D. in geological sciences at Columbia University in 1993. He has continued his investigations of seafloor topography as a research scientist, first at Columbia's Lamont-Doherty Earth Observatory and then at the Institute of Arctic and Alpine Research at the University of Colorado at Boulder. Since 1998 Pratson has been associate professor of sedimentary geology at Duke University's Nicholas School of the Environment and Earth Sciences. Haxby earned his doctorate from Cornell University in 1978. Since then, he has conducted studies of the ocean basins as a research scientist at Lamont-Doherty Earth Observatory.

THREE TYPES OF MARGINS exist along the borders of North America. The sides facing the Atlantic Ocean and the Gulf of Mexico are termed passive margins, because they are embedded within the North American Plate and simply ride along with this great slab as it moves. The western margin of the U.S., on the other hand, lies along the leading edge of the North American Plate, where it bumps and grinds its way past oceanic crust underlying the Pacific Ocean. This collision takes two forms. Through most of the length of California (a strike-slip margin), the North American Plate slips sideways past the Pacific Plate along a system of near-vertical fractures in the earth collectively known as the San Andreas Fault. Farther up the coast (a convergent margin), the North American Plate is bulldozing its way over a sliver of oceanic crust named the Juan de Fuca Plate.

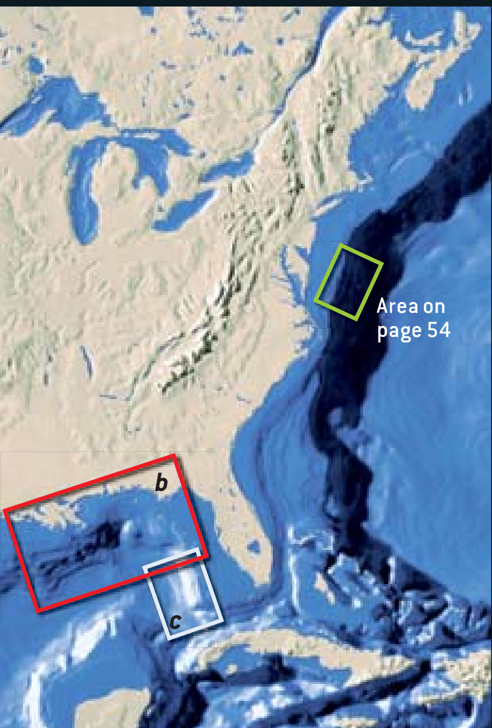


EDWARD BELL

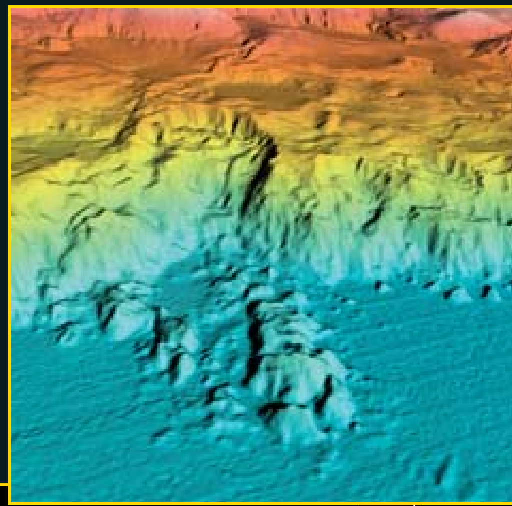
CRATERLIKE FEATURES pockmark the seafloor offshore from Mississippi to eastern Texas. These small basins are filled with thick accumulations of sediment—and, in some spots, billions of barrels of oil and gas. The ridges and domes between the basins hide shallowly buried salt bodies of various sizes and



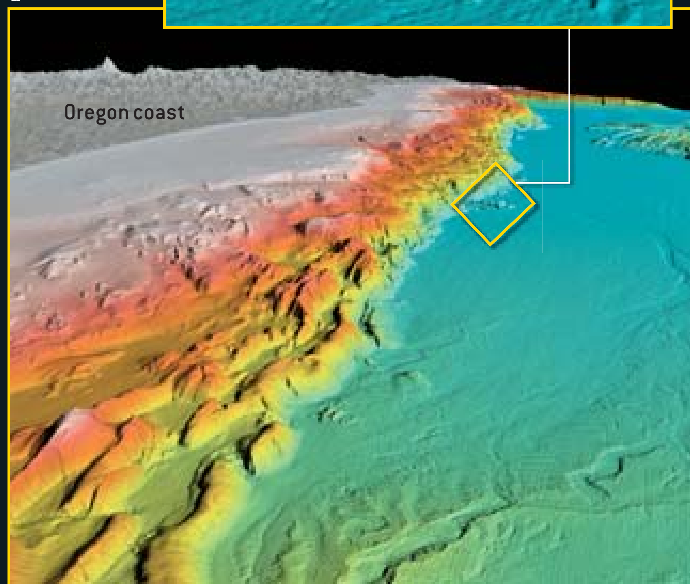
ILLUSTRATIONS BY LINCOLN F. PRATSON AND WILLIAM F. HAXBY, EXCEPT AS NOTED



SEAFLOOR FAILURE carved a pocket in the continental slope offshore from central Oregon. The trail of debris extending outward from the crescent-shaped embayment in the continental slope marks the path the material traveled. The excavated pocket is six kilometers wide—about the width of the island of Manhattan. Some of the dislodged blocks that now rest on the continental rise are as tall as a modern skyscraper. The collapse that sent these huge chunks tumbling down the slope was most likely triggered by an earthquake. Such dramatic failures of the continental slope can generate violent tsunamis that may inundate the coast nearby or travel across the Pacific and create havoc on distant shores.

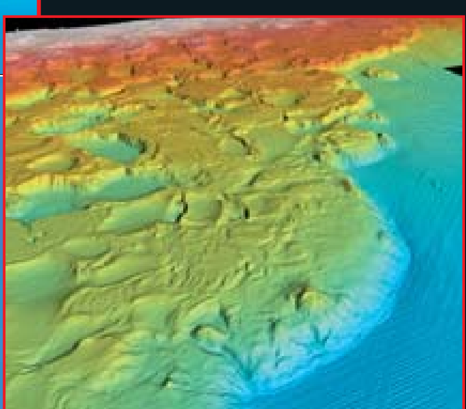


a

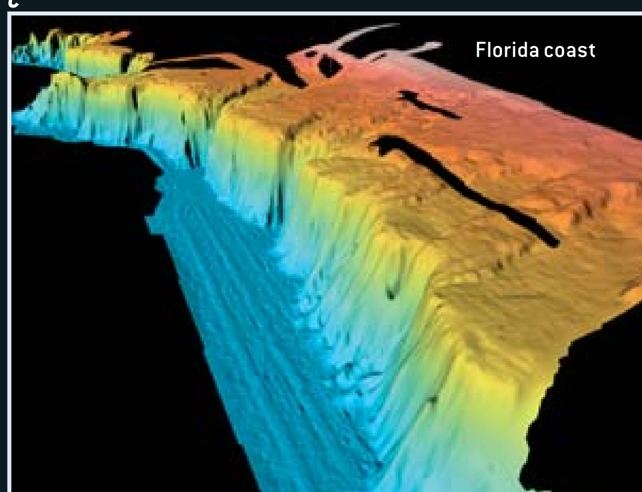


FOLDED CARPET of sediments covers the seafloor offshore of Oregon. The undulations result from the head-on collision between the North American and Juan de Fuca plates. Like a colossal bulldozer, the North American Plate scrapes sediments off the down-going Juan de Fuca Plate and piles them into folds. To the north (*lower left*), the folds of sediment form distinct ridges. To the south (*upper right*), where part of the Juan de Fuca Plate breaks through its sedimentary cover, the folds are stacked so closely that they form terraces.

shapes. This salt formed initially by the evaporation of Gulf of Mexico waters (some 180 million years ago). The salt layer was then buried by a massive load of sediment eroded from the Rocky Mountains and carried into the gulf by the Mississippi River. Because salt resists compression, it will flow rather than compact. Hence, under the weight of the overlying sediment, giant blebs of the salt have bubbled upward and spread out toward the open sea.



c



MILE-HIGH CLIFF marks the edge of the continental slope west of Florida. This undersea precipice, known as the Florida Escarpment, stands more than four times as high as the Empire State Building. Whereas the tilt of the continental slope elsewhere is typically just a few degrees, the face of the escarpment is, on average, slanted at 35 degrees. In many places the walls of the escarpment are near vertical. The seafloor here is made up of the countless skeletons of marine organisms that have cemented together. The gradual accumulation of this material once formed a gently dipping ramp. But some force, perhaps great sweeping currents, eroded the base of the slope. Today extremely salty groundwaters seep out of the face of the escarpment and dissolve the rock there. Weakened by this decay, the slope can collapse, taking a good deal of overlying material with it. Curiously, little if any vestige of the vast amount of material worn away can be found along the base of the cliff.

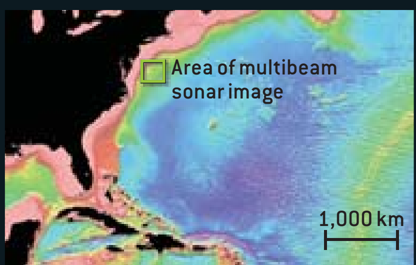
SEAFLOOR MAPPING TOOLS

OUR MULTIBEAM sonar images represent just one way scientists can visualize the seafloor. Other approaches are also used, and each has its peculiar advantages and shortcomings.

Satellites (*a*) cannot measure seafloor depth directly, but they can sense variations in the elevation of the water at the surface of the ocean. The U.S. Navy's Geosat satellite, for example, can measure the distance to the ocean surface to within five centimeters by bouncing radar pulses off the water below it. Because the precise position of the satellite is known, such determinations provide a measure of sea-surface height.

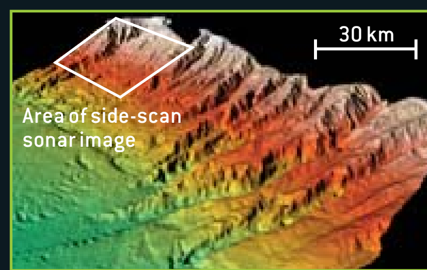
The ocean surface can vary in relief by as much as 200 meters. These undulations reflect minute differences in Earth's gravity from place to place that cause water to distribute itself unevenly. Most commonly, these gravitationally induced variations in the ocean surface are caused by rugged seafloor topography. For instance, a massive, submerged volcano that is 2,000 meters tall and 40 kilometers wide will pull water toward it, producing a bulge about two meters high in the ocean surface above it. But undersea features smaller than 10 kilometers across do not generally possess sufficient mass to affect the ocean surface and thus go undetected by satellite radars. What is more, gravity variations (particularly near continental margins) can reflect differences in the density of the underlying rock rather than topography. Still, satellites provide broad, if less than perfect, maps of regions not yet surveyed with ships.

Multibeam sonar (*b*) bounces sound off the seafloor to gauge ocean depth. In contrast to simple echo sounders, this modern technique employs an array of sound sources and listening devices mounted on the hull of the survey vessel. Every few seconds the sources emit a burst that reaches only a slim strip of seafloor aligned perpendicularly to the direction the ship is moving. At the same time, the listening devices begin recording sounds reflected from the bottom. This equipment is arranged to detect sounds coming only from within a series of narrow seafloor corridors that are aligned parallel to the ship's direction. Thus, the sound reflections received



WALTER H.F. SMITH *NOAA*

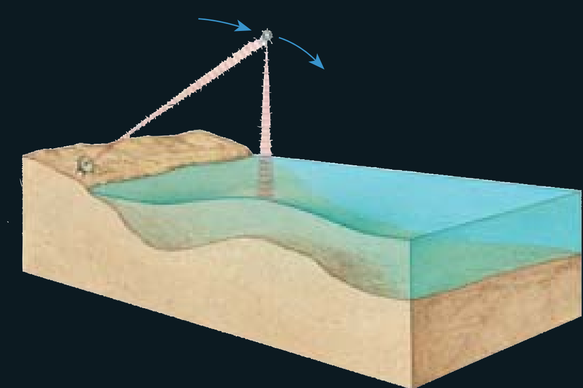
SATELLITE RADAR IMAGE



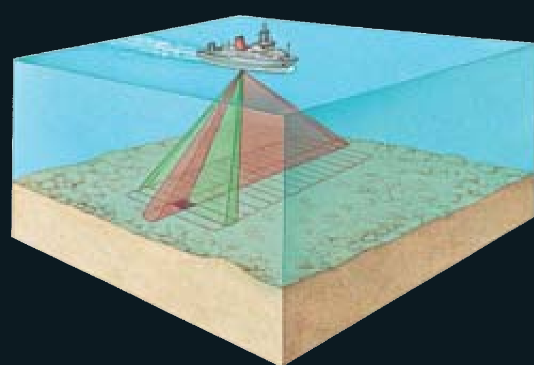
L. F. PRATSON AND W. F. HAXBY

MULTIBEAM SONAR IMAGE

a



b



ROBERTO OSTI

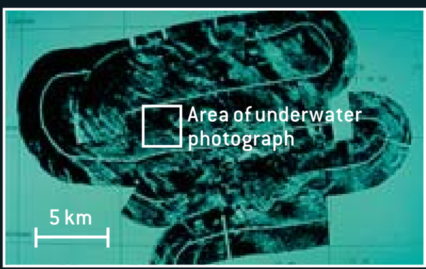
maritime nations declared the waters and seafloor within 200 miles of their shores to be "Exclusive Economic Zones." To help assess the value of the vast undersea expanse claimed by the U.S., the National Oceanic and Atmospheric Administration began surveying parts of the newly annexed area in 1983. That effort (which continued until 1993) mapped more than 200,000 square kilometers of the seafloor off the coasts of the Atlantic and Pacific oceans and the Gulf of Mexico.

Over this same period, the National Science Foundation funded two smaller surveys to study parts of the seafloor near the coasts of New Jersey and western Florida. All the vessels involved used multibeam sonars, the most modern form of instrumentation available for measuring the topography of the ocean bottom.

These surveys provide unprecedented views of the country's continental slope. Although no sunlight actually penetrates to these great depths, computers can render images of seafloor vistas as they would appear with the oceans drained. Such a perspective could be particularly valuable in planning industrial activities offshore. For example, submarine cables increasingly carry international communications, and petroleum producers are moving drilling platforms into ever greater depths of water. These enterprises require maps of where the seafloor appears to be stable—not prone to subsea avalanches or violent currents. Disposal of at the ship emanate from the regions where the slim strip of sound and the listening corridors intersect. The timing of these reflections provides a profile of seafloor depth. Such profiles are recorded every few seconds while the survey ship moves over the seafloor, and so successive observations build up a continuous swath of coverage along the ship's track. By running the ship in the same pattern one mows a lawn, scientists can produce a complete map of an area. With less than 200 vessels outfitted with the necessary equipment, however, charting the entire seafloor in this way would require hundreds of years.

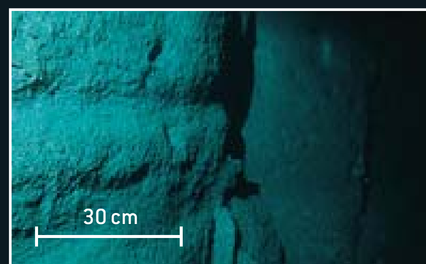
Side-scan sonar (*c*) provides yet a different perspective on what the seafloor looks like. The equipment is usually attached to a "sled" that is towed behind a ship. Two sonar units, affixed to either side of the sled, act as both sound sources and listening devices. These units emit bursts of sound outward, to either side. If the seafloor is flat and smooth, none of the energy emitted will be reflected back (as with a beam of light directed obliquely onto a mirror). But if the seafloor is rough, the sound hitting the bottom will be scattered in all directions, and some will return to the sonar sled (just as a beam of light illuminating ground glass will reflect in all directions). By equating the amplitude of the recorded echoes to different shades of gray and displaying the results to show the distance from the sled, scientists can obtain an image of the texture of the seafloor that looks similar to a black-and-white photograph. But like a single aerial photo, a side-scan sonar image does not indicate the heights of the surface below.

The most accurate and detailed view of the seafloor is provided by underwater photography (*d*), using either cameras towed along the bottom, piloted submersibles or remotely operated vehicles. Such camera-carrying equipment gives researchers the opportunity to explore the seafloor up close. Yet because even the most intense illumination does not penetrate seawater effectively, photographic views obtained in this way are limited to the short distances that artificial beams of light can penetrate.
—L.F.P. and W.F.H.



L. F. PRATSON AND W.B.F. RYAN

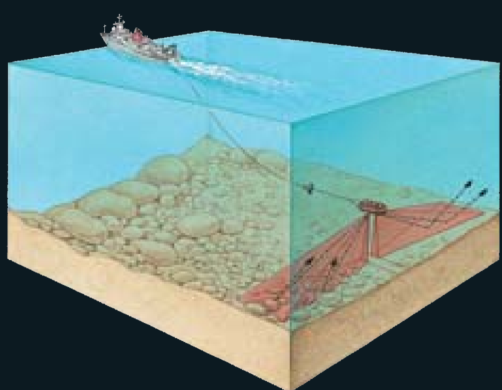
SIDE-SCAN SONAR IMAGE



LINCOLN F. PRATSON

UNDERWATER PHOTOGRAPH

c



d

ROBERTO OSTI

waste at sea also demands this information, because currents running along the bottom can disturb the sites where waste settles. Bottom surveys further help geologists to locate offshore fault systems and to assess their risk of triggering earthquakes.

On a broader scientific level, undersea mapping is providing fundamental knowledge about the geologic forces that shape the ocean floor. Images such as those we have created offer scientists a way to take in vast stretches of undersea terrain in a glance—an ability they have long enjoyed while studying the surface of distant moons and planets. That perspective now offers some fascinating new insights into the marvelously complex evolution of Earth. SA

MORE TO EXPLORE

Imaging the Ocean Floor: History and State of the Art. Peter R. Vogt and Brian E. Tucholke in *Geology of North America*, Vol. M. Edited by P. R. Vogt and B. E. Tucholke. Geological Society of America, 1986.

What Is the Slope of the U.S. Continental Slope? Lincoln F. Pratson and William F. Haxby in *Geology*, Vol. 24, No. 1, pages 3–6; January 1996.

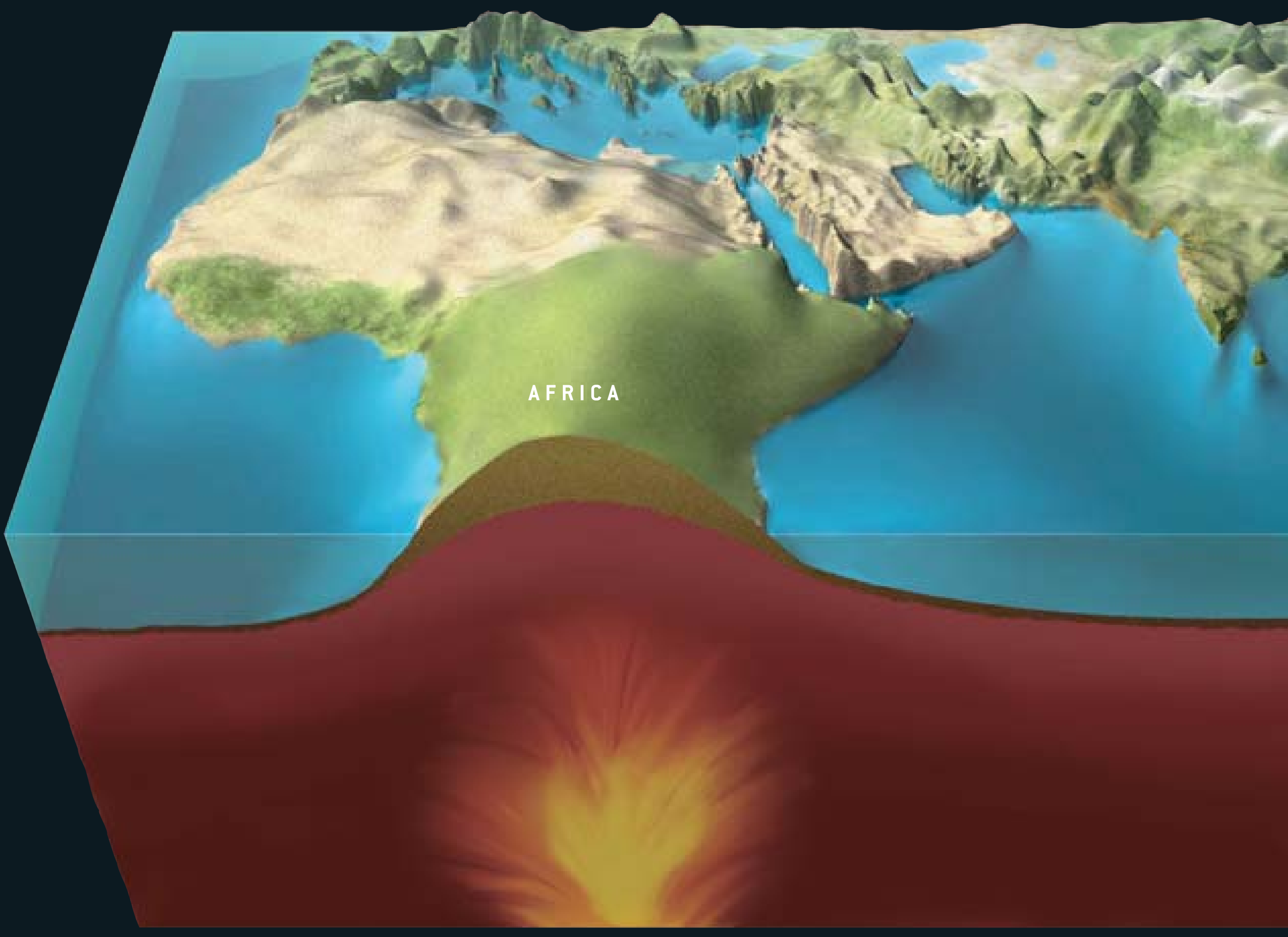
Global Sea Floor Topography from Satellite Altimetry and Ship Depth Soundings. W. H. Smith et al. in *Science*, Vol. 277, pages 1956–1962; September 26, 1997.

Categorizing the Morphologic Variability of Siliciclastic Passive Continental Margins. D. O'Grady, J. Syvitski, L. F. Pratson and J. F. Sarg in *Geology*, Vol. 28, No. 3, pages 207–210; March 2000.

NOAA's National Geophysical Data Center Web site is available at **www.ngdc.noaa.gov/ngdc.html**

Sculpting Earth from Inside Out

By Michael Gurnis



Powerful motions deep inside the planet do not merely shove fragments of the rocky shell horizontally around the globe— they also lift and lower entire continents

ENIGMATIC DIPS AND SWELLS have occurred over continent-size swaths of Earth's surface several times in the past. Southern Africa has been lifted about 300 meters over the past 20 million years, for example, and a sunken continent's highest peaks today form the islands of Indonesia.

ALL ILLUSTRATIONS BY DAVID FIERSTEIN

Credit for sculpting Earth's surface typically goes to violent collisions between tectonic plates, the mobile fragments of the planet's rocky outer shell. The mighty Himalayas shot up when India rammed into Asia, for instance, and the Andes grew as the Pacific Ocean floor plunged beneath South America. But even the awesome power of plate tectonics cannot fully explain some of the planet's most massive surface features.

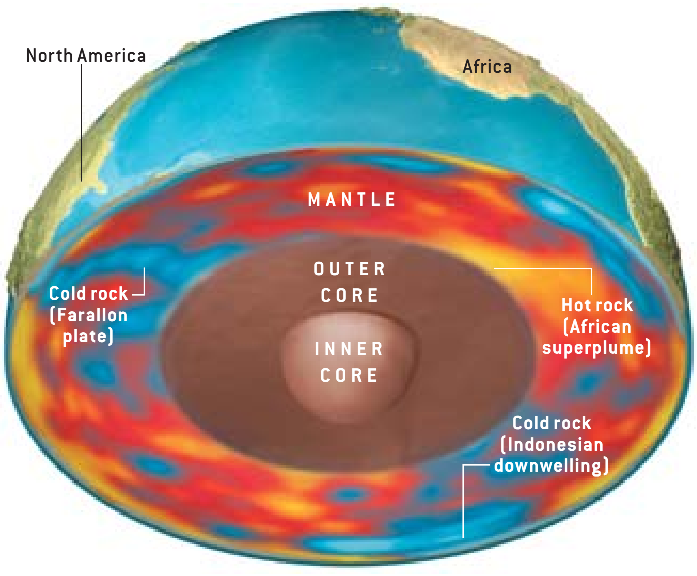
Take southern Africa, which boasts one of the world's most expansive plateaus, more than 1,500 kilometers across and almost two kilometers high. Geologic evidence shows that southern Africa, and the surrounding ocean floor, has been rising slowly for the past 100 million years, even though it has not experienced a tectonic collision for nearly 400 million years.

The African superswell, as this uplifted landmass is known, is just one example of dramatic vertical movement by a broad chunk of Earth's surface. In other cases from the distant past, vast stretches of Australia and North America bowed down hundreds of meters—and then popped up again.

Scientists who specialize in studying Earth's interior have long suspected that activity deep inside Earth was behind such vertical changes at the surface. These geophysicists began searching for clues in the mantle—the middle layer of the planet. This region of scalding-hot rock lies just below the jigsaw configuration of tectonic plates and extends down more than 2,900 kilometers to the outer edge of the globe's iron core. Researchers learned that variations in the mantle's intense heat and pressure enable the solid rock to creep molasseslike over thousands of years. But they could not initially decipher how it could give rise to large vertical motions. Now, however, powerful computer models that combine snapshots of the mantle today with clues about how it might have behaved in the past are beginning to explain why parts of Earth's surface have undergone these astonishing ups and downs.

The mystery of the African superswell was among the easiest to decipher. Since the early half of the 20th century, geophysicists have understood that over the immense expanse of geologic time, the mantle not only creeps, it churns and roils like a pot of thick soup about to boil. The relatively low density of the hottest rock makes that material buoyant, so it slowly ascends; in contrast, colder, denser rock sinks until heat escaping the molten core warms it enough to make it rise again. These three-dimensional motions, called convection, are known to enable the horizontal movement of tectonic plates, but it seemed unlikely that the forces they created could lift and lower the planet's surface. That skepticism about the might of the mantle began to fade away when researchers created the first blurry images of Earth's interior.

About 20 years ago scientists came up with a way to make three-dimensional snapshots of the mantle by measuring vibrations that are set in motion by earthquakes originating in the planet's outer shell. The velocities of these vibrations, or seismic waves, are determined by the chemical composition, temperature and pressure of the rocks they travel through. Waves become sluggish in hot, low-density rock, and they speed up in colder, denser regions. By recording the time it takes for seismic waves to travel from an earthquake's epicenter to a particular recording station at the surface, scientists can infer the temperatures and densities in a given segment of the interior. And by compiling a map of seismic velocities from thousands of earthquakes around the globe, they can begin to map temperatures and densities throughout the mantle.



MANTLE MAP integrates measurements of thousands of earthquake vibrations, or seismic waves, that have traveled through the planet. Regions where waves moved quickly (*blue*) usually denote cold, dense rock. Regions where waves slowed down (*yellow*) denote hot, less compact rock. Under southern Africa and the South Atlantic lies a pocket of sluggish velocities—a buoyant blob of hot rock called the African superplume. The map also reveals cold, sinking material that is tugging on North America and Indonesia.

SOURCE: JEROEN RITSEMA *California Institute of Technology*

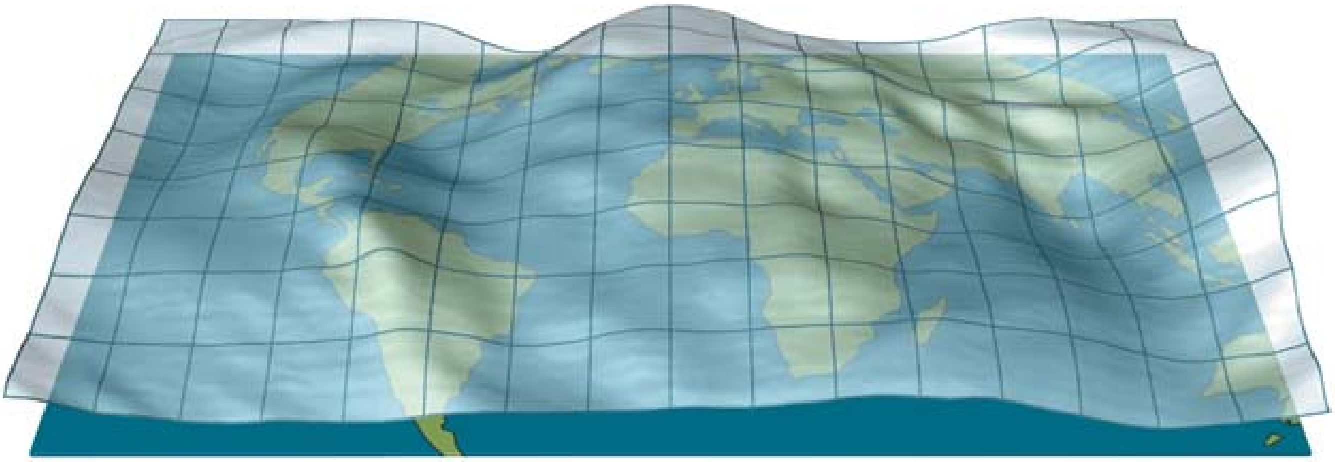
These seismic snapshots, which become increasingly more detailed as researchers find more accurate ways to compile their measurements, have recently revealed some unexpectedly immense formations in the deepest parts of the mantle. The largest single structure turns out to lie directly below Africa's southern tip. About five years ago seismologists Jeroen Ritsema, now at the Paris Geophysical Institute (IPGP), and Hendrik-Jan van Heijst, now at Shell Research, calculated that this mushroom-shaped mass stretches some 1,400 kilometers upward from the core and spreads across several thousand kilometers [*see illustration below*].

The researchers immediately began to wonder whether this enormous plume could be shoving Africa skyward. Because the plume is a region where seismic waves are sluggish, they assumed that it was hotter than the surrounding mantle. The basic physics of convection suggested that a hot plume was likely to be rising. But a seismic snapshot records only a single moment in time and thus only one position of a structure. If the plume were of a different composition than the surrounding rock, for instance, it could be hotter and still not rise. So another geophysicist, Jerry X. Mitrovica of the University of Toronto, and I decided to create a time-lapse picture of what might be happening. We plugged the plume's shape and estimated density, along with estimates of when southern Africa began rising, into a computer program that simulates mantle convection. By doing so, we found that the plume is indeed buoyant enough to rise slowly within the mantle and strong enough—a veritable superplume—to push Africa upward as it goes [see "The Mid-Cretaceous Superplume Episode," on page 22].

But whether the African superplume is merely hot mantle or actually has a different composition remains a topic of active research. Two years ago seismologists Donald V. Helmberger and Sidao Ni teamed up with geodynamists Eh Tan and me, all at the California Institute of Technology, to address this question. By monitoring earthquake-induced seismic waves that grazed the eastern side of the plume, we discovered that the plume had sharp sides. Computer models that matched this sharpness, and the observation that it was tilted from the vertical, suggested that a large, chemically distinct thermal anomaly was actively upwelling within the mantle.

Seismic snapshots and computer models—the basic tools of geophysicists—were enough to solve the puzzle of the African superswell, but resolving the up-and-down movements of North America and Australia was more complicated and so was accomplished in a more circuitous way. Geophysicists who think only about what the mantle looks like today cannot fully explain how it sculpts Earth's surface. They must therefore borrow from the historical perspective of traditional geologists who think about the way the surface has changed over time.

The insights that would help account for the bobbings of



BULGES AND TROUGHS in the transparent surface above the world map represent natural variations in Earth's gravitational field. High points indicate stronger-than-normal gravity caused by a pocket of excess mass within the planet's interior; low areas occur above regions where a deficiency of mass produces a band of low gravity. Such differences in gravity hint at the location of oddities in the structure of Earth's mantle.

Australia and North America began to emerge with investigations of a seemingly unrelated topic: the influence of mantle density on Earth's gravitational field. The basic principles of physics led scientists in the 1960s to expect that gravity would be lowest above pockets of hot rock, which are less dense and thus have less mass. But when geophysicists first mapped Earth's gravitational variations, they found no evidence that gravity correlated with the cold and hot parts of the mantle—at least not in the expected fashion.

Indeed, in the late 1970s and early 1980s Clement G. Chase uncovered the opposite pattern. When Chase, now at the University of Arizona, considered geographic scales of more than 1,500 kilometers, he found that the pull of gravity is strongest not over cold mantle but over isolated volcanic regions called hot spots. Perhaps even more surprising was what Chase noticed about the position of a long band of low gravity that passes from Hudson Bay in Canada northward over the North Pole, across Siberia and India, and down into Antarctica. Relying on estimates of the ancient configuration of tectonic plates, he showed that this band of low gravity marked the location of a series of subduction zones—that is, the zones where tectonic plates carrying fragments of the seafloor plunge back into the mantle—from 125 million years ago. The ghosts of ancient subduction zones seemed to be diminishing the pull of gravity. But if cold, dense chunks of seafloor were still sinking through the mantle, it seemed that gravity would be high above these spots, not low, as Chase observed.

In the mid-1980s geophysicist Bradford H. Hager, now at the Massachusetts Institute of Technology, resolved this apparent paradox by proposing that factors other than temperature might create pockets of extra or deficient mass within the mantle. Hager developed his theory from the physics that describes moving fluids, whose behavior the mantle imitates over the long term. When a low-density fluid rises upward, as do the hottest parts of the mantle, the force of the flow pushes up the higher-density fluid above it. This gentle rise atop the upwelling itself creates an excess of mass (and hence stronger gravity) near the planet's surface. By the same token, gravity can be lower over cold, dense material: as this heavy matter sinks, it drags down mass that was once near the surface. This theory explained why the ghosts of subduction zones could generate a band of low gravity: some of that cold, subducted seafloor must still be sinking within the mantle—and towing the planet's surface downward in the process. If Hager's explanation was correct, it meant that the mantle did not merely creep horizontally near the planet's surface; whole segments of the mantle's up-and-down movements also reached the surface and altered it significantly. Areas that surged upward would push the land above it skyward, and areas that sank would drag down the overlying continents as they descended.

BOBBING CONTINENTS

AT THE SAME TIME that Chase and Hager were discovering a mechanism that could dramatically lift and lower Earth's surface, geologists were beginning to see evidence that continents might actually have experienced such dips and swells in the past. Geologic formations worldwide contain evidence that sea level fluctuates over time. Many geologists suspected that this fluctuation would affect all continents in the same way, but a few of them advanced convincing evidence that the most momentous changes in sea level stemmed from vertical motions of continents. As one continent moved, say, upward relative to other landmasses, the ocean surface around that continent would become lower while sea level around other landmasses would stay the same.

Most geologists, though, doubted the controversial notion that continents could move vertically—even when the

THE AUTHOR

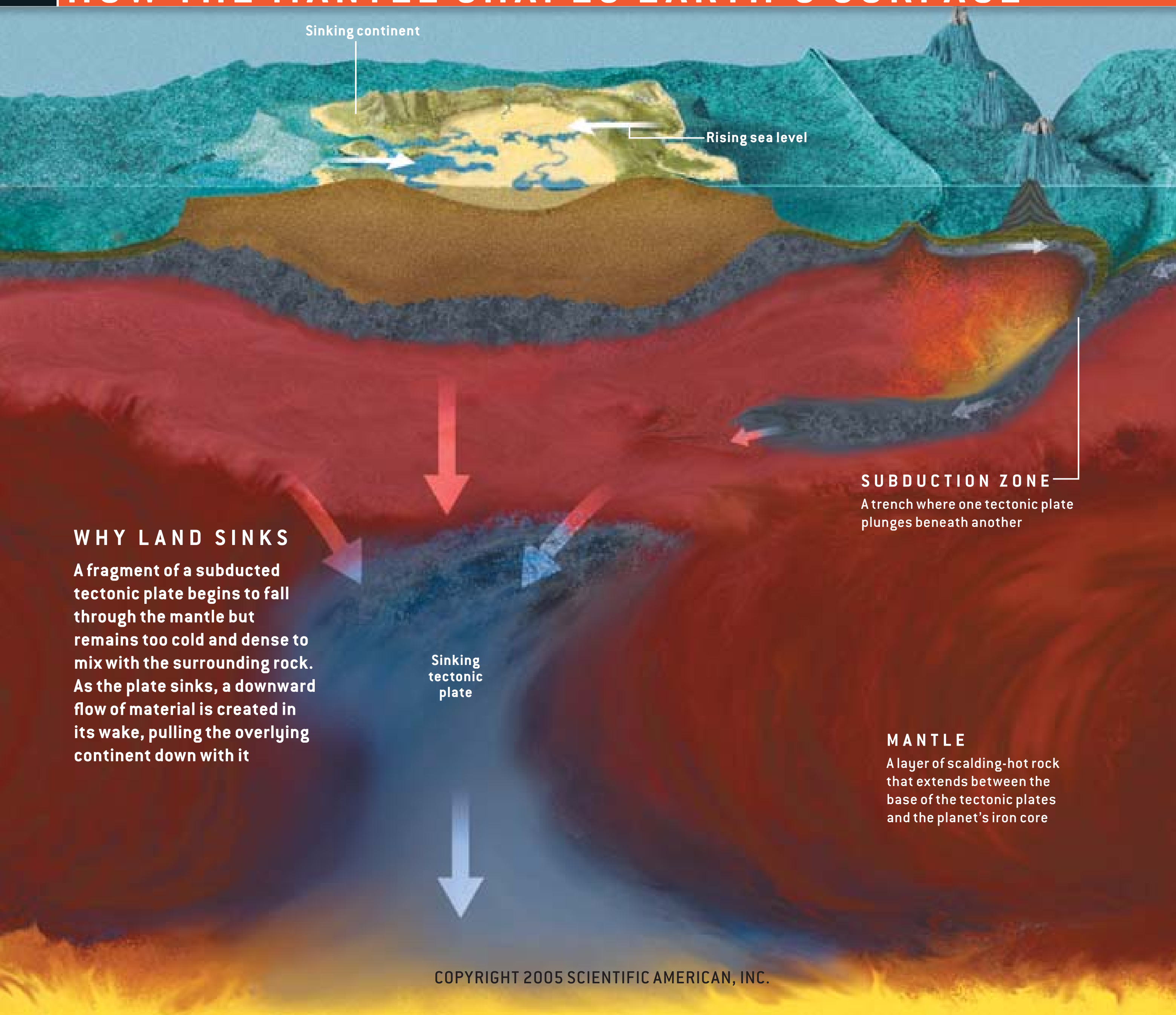
MICHAEL GURNIS is a geophysicist who is interested in the dynamics of plate tectonics and Earth's interior. These physical processes, which govern the history of the planet, have intrigued him since he began studying geology as an undergraduate 25 years ago. With his research group at the California Institute of Technology, Gurnis now develops computer programs that simulate the evolving motions of the mantle and reveal how those motions have shaped the planet over time. Gurnis also leads a national effort to develop the next generation of simulation software for geophysics research in the U.S.

first indications of the bizarre bobbing of Australia turned up in the early 1970s. Geologist John J. Veevers of Macquarie University in Sydney examined outcrops of ancient rock in eastern Australia and discovered that sometime in the Early Cretaceous period (about 130 million years ago), a shallow sea rapidly covered that half of Australia while other continents flooded at a much more leisurely pace. Sea level climaxed around those landmasses by the Late Cretaceous (about 70 million years ago), but by then the oceans were already retreating from Australia's shores. The eastern half of the continent must have sunk several hundred meters relative to other landmasses and then popped back up before global sea level began to fall.

Veevers's view of a bobbing continent turned out to be only part of Australia's enigmatic story. In 1978 geologist Gerard C. Bond, now at Columbia University's Lamont-Doherty Earth Observatory, discovered an even stranger turn of events while he was searching global history for examples of vertical continental motion. After Australia's dip and rise during the Cretaceous, it sank again, this time by 180 meters, between then and the present day. No reasonable interpretation based on plate tectonics alone could explain the widespread vertical motions that Bond and Veevers uncovered. Finding a satisfactory explanation would require scientists to link this information with another important clue: Hager's theory about how the mantle can change the shape of the planet's surface.

The first significant step in bringing these clues together was the close examination of another up-and-down example from Bond's global survey. In the late 1980s this work inspired Christopher Beaumont, a geologist at Dalhousie University in

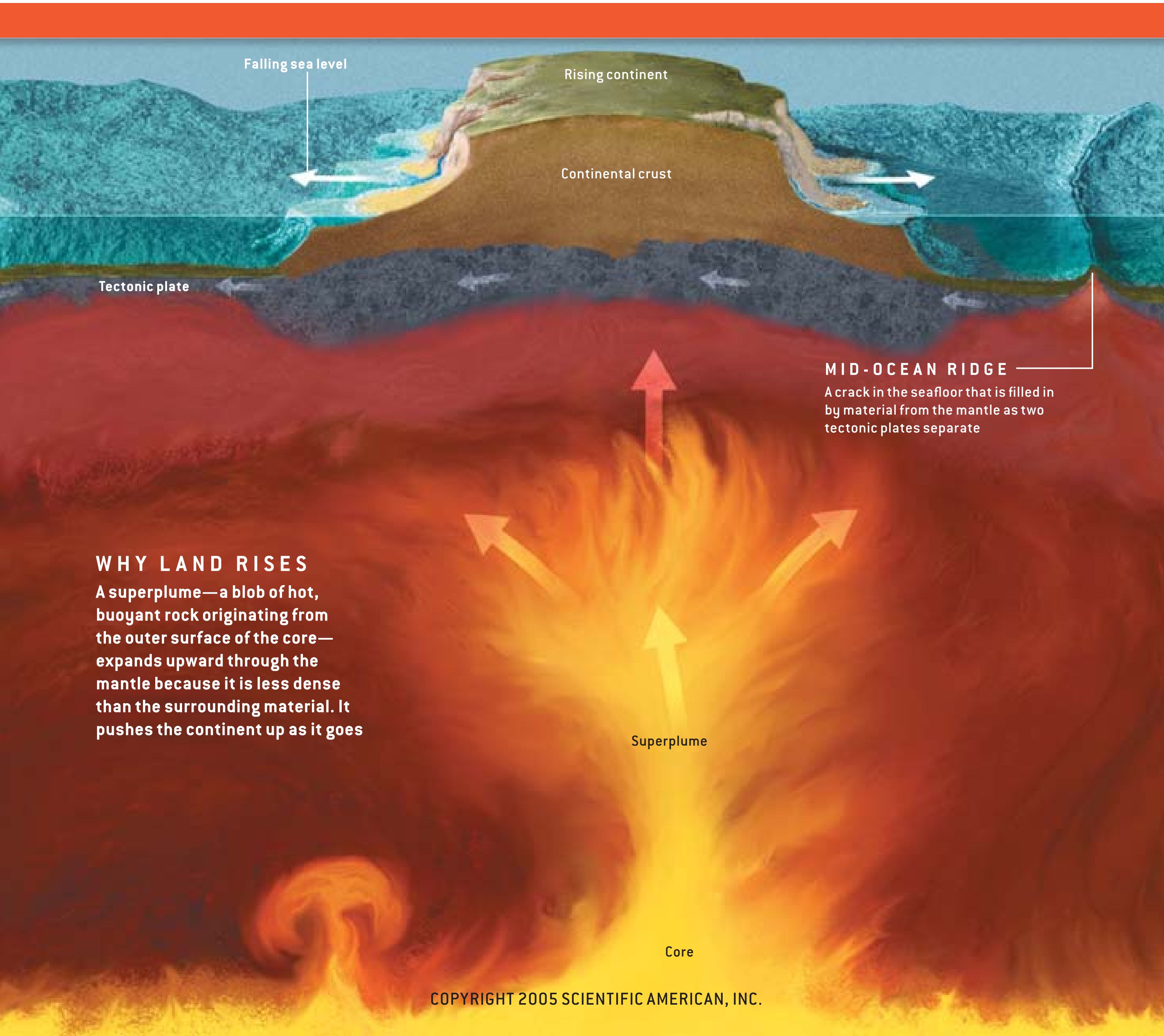
HOW THE MANTLE SHAPES EARTH'S SURFACE



Halifax, Nova Scotia, to tackle a baffling observation about Denver. Although the city's elevation is more than 1,600 meters above sea level, it sits atop flat, undeformed marine rocks created from sediments deposited on the floor of a shallow sea during the Cretaceous period. Vast seas covered much of the continents during that time, but sea level was no more than about 120 meters higher than it is today. This means that the ocean could never have reached Denver's current position—unless this land was first pulled down several hundred meters to allow waters to flood inland.

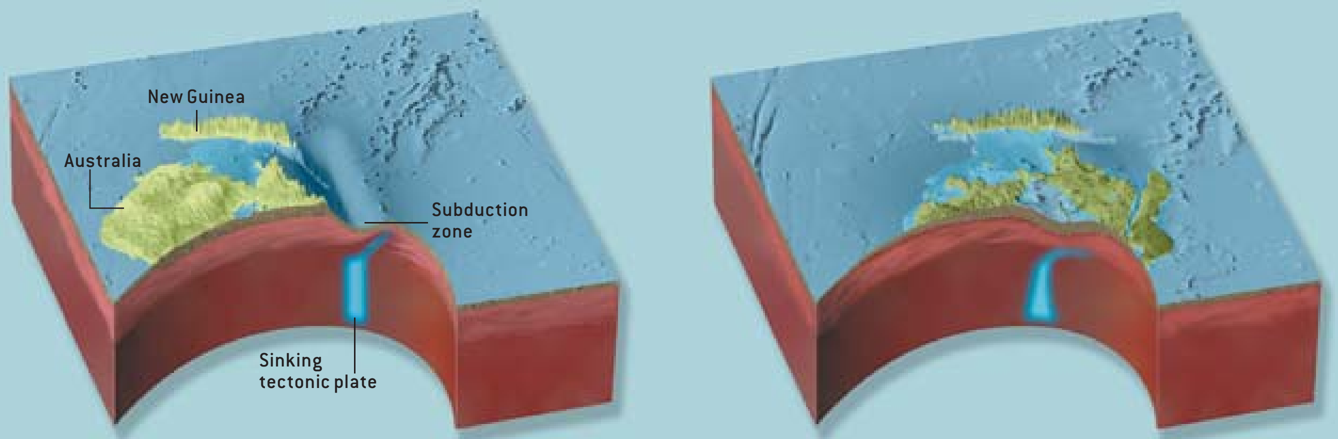
Based on the position of North America's coastlines during the Cretaceous, Beaumont estimated that this bowing downward and subsequent uplift to today's elevation must have affected an area about 1,000 kilometers across. This geographic scale was problematic for the prevailing view that plate tectonics alone molded the surface. The mechanism of plate tectonics permits vertical motions within only 200 kilometers or so of plate edges, which are thin enough to bend like a stiff fishing pole when forces act on them. But the motion of North America's interior happened several hundred kilometers inland—far from the influence of plate collisions. An entirely different mechanism had to be operating.

Beaumont knew that subducted slabs of ancient seafloor might sit in the mantle below North America and that such slabs could theoretically drag down the center of a continent. To determine whether downward flow of the mantle could have caused the dip near Denver, Beaumont teamed up with Mitrovica, then a graduate student at the University of Toronto, and Gary T. Jarvis of York University in Toronto. They found that the sinking of North America during the Creta-



AUSTRALIA'S UPS AND DOWNS

A computer model reveals how the ghost of an ancient subduction zone dragged down a continent.



130 MILLION YEARS AGO

Australia is bordered by a subduction zone, a deep trench where the tectonic plate to the east plunges into the mantle. The sinking plate (*blue*) pulls the surrounding mantle and the eastern edge of Australia down with it. Later, subduction ceases, and the continent begins to drift eastward.

90 MILLION YEARS AGO

The entire eastern half of Australia sinks about 300 meters below sea level as the continent passes eastward over the sinking tectonic plate. About 20 million years later the plate's downward pull diminishes as it descends into the deeper mantle. As a result, the continent then pops up again.

ceous could have been caused by a plate called the Farallon as it plunged into the mantle beneath the western coast of North America. Basing their conclusion on a computer model, the research team argued that the ancient plate thrust into the mantle nearly horizontally. As the plate began sinking, it created a downward flow in its wake that tugged North America low enough to allow the ocean to rush in. As the Farallon plate sank deeper, the power of its trailing wake decreased. The continent's tendency to float eventually won out, and North America resurfaced.

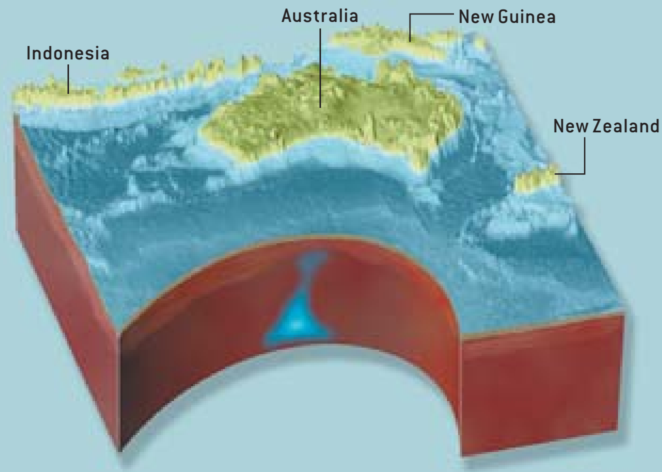
When the Canadian researchers advanced their theory in 1989, the Farallon plate had long since vanished into the mantle, so its existence had only been inferred from geologic indications on the bottom of the Pacific Ocean. At that time, no seismic images were of high enough resolution to delineate a structure as small as a sinking fragment of the seafloor. Then, in 1996, new images of the mantle changed everything. Stephen P. Grand of the University of Texas at Austin and Robert D. van der Hilst of M.I.T., seismologists from separate research groups, presented two images based on entirely different sets of seismic measurements. Both pictures showed virtually identical structures, especially the cold-mantle downwellings associated with sinking slabs of seafloor. The long-lost Farallon plate was prominent in the images as an arching slab 1,500 kilometers below the eastern coast of the U.S.

MOVING DOWN UNDER

CONNECTING THE BOBBING MOTION of North America to the subduction of the seafloor forged a convincing link between ancient sea-level change and goings-on in the mantle. It also became clear that the ancient Farallon slab sits within the band of low gravity that Chase had observed two decades earlier. I suspected that these ideas could also be applied to the most enigmatic of the continental bobbings, that of Australia during and since the Cretaceous. I had been simulating mantle convection with computer models for 15 years, and many of my results showed that the mantle was in fact able to lift the surface by hundreds of meters—a difference easily great enough to cause an apparent drop in sea level. Like Chase, Veevers and other researchers before me, I looked at the known history of plate tectonics for clues about whether something in the mantle could have accounted for Australia's bouncing. During the Cretaceous period, Australia, South America, Africa, India, Antarctica and New Zealand were assembled into a vast supercontinent called Gondwana, which had existed for more than 400 million years before it fragmented into today's familiar landmasses. Surrounding Gondwana for most of this time was a huge subduction zone where cold oceanic plates plunged into the mantle.

I thought that somehow the subduction zone that surrounded Gondwana for hundreds of millions of years might have caused Australia's ups and downs. I became more convinced when I sketched the old subduction zones on maps of ancient plate configurations constructed by R. Dietmar Müller, a seagoing geophysicist at Sydney University. The sketches seemed to explain the Australian oddities. Australia would have passed directly over Gondwana's old subduction zone at the time it sank.

To understand how the cold slab would behave in the mantle as Gondwana broke apart over millions of years, Müller and I joined Louis Moresi, now at Monash Univer-



PRESENT DAY

Australia lies north of its former site, having been pushed there by activity in adjacent tectonic plates beginning about 45 million years ago. The entire continent has dropped relative to its greatest elevation as the result of a downward tug in the mantle under Indonesia—a landmass that is also sinking.

sity in Australia, to run a computer simulation depicting the mantle's influence on Australia over time. We knew the original position of the ancient subduction zone, the history of horizontal plate motions in the region and the estimated properties—such as viscosity—of the mantle below. Operating under these constraints, the computer played out a scenario for Australia that fit our hypotheses nearly perfectly [*see box above*].

The computer model started 130 million years ago with ocean floor thrusting into the mantle beneath eastern Australia. As Australia broke away from Gondwana, it passed over the cold, sinking slab, which sucked the Australian plate downward. The continent rose up again as it continued its eastward migration away from the slab.

Our model resolved the enigma of Australia's motion during the Cretaceous, originally observed by Veevers, but we were still puzzled by the later continentwide sinking of Australia that Bond discovered. With the help of another geophysicist, Carolina Lithgow-Bertelloni, now at the University of Michigan at Ann Arbor, we confirmed Bond's observation that as Australia moved northward toward Indonesia after the Cretaceous, it subsided by about 200 meters. Lithgow-Bertelloni's global model of the mantle, which incorporated the history of subduction, suggested that Indonesia is sucked down more than any other region in the world because it lies at the intersection of enormous, present-day subduction systems in the Pacific and Indian oceans. And as Indonesia sinks, it pulls Australia down with it. Today Indonesia is a vast submerged continent—only its highest mountain peaks protrude above sea level.

Which brings us back to Africa. In a sense, Indonesia and Africa are opposites: Indonesia is being pulled down while Africa is being pushed up. These and other changes in the mantle that have unfolded over the past few hundred million years are intimately related to the supercontinent Gondwana. The huge band of low gravity that Chase discovered 30 years ago is created by the still-sinking plates of a giant subduction zone that once encircled the vast southern landmass. At the center of Gondwana was southern Africa, which means that the mantle below this region was isolated from the chilling effects of sinking tectonic plates at that time—and for the millions of years since. This long-term lack of cold, downward motion below southern Africa explains why a hot superplume is now erupting in the deep mantle there.

With all these discoveries, a vivid, dynamic picture of the motions of the mantle has come into focus. Researchers are beginning to see that these motions sculpt the surface in more ways than one. They help to drive the horizontal movement of tectonic plates, but they also lift and lower the continents. Perhaps the most intriguing discovery is that motion in the deep mantle lags behind the horizontal movement of tectonic plates. Positions of ancient plate boundaries can still have an effect on the way the surface is shaped many millions of years later.

Our ability to view the dynamics of mantle convection and plate tectonics will rapidly expand as new ways of observing the mantle and techniques for simulating its motion are introduced. Higher-resolution seismic images will also play a pivotal role in revealing what the mantle looks like today. A program for obtaining such higher-resolution images began last year. Called the USArray project, it deploys 400 seismometers that rove across the country with the aim of providing an 80-kilometer-resolution view into the upper 1,300 kilometers of the mantle below the U.S.

Plans to make unprecedented images and measurements of the mantle in the coming decade, together with the use of ever more powerful supercomputers, foretell an exceptionally bright future for deciphering the dynamics of Earth's interior. Already, by considering the largest region of the planet—the mantle—as a chunk of rock with a geologic history, earth scientists have made extraordinary leaps in understanding the ultimate causes of geologic changes at the surface. SA

MORE TO EXPLORE

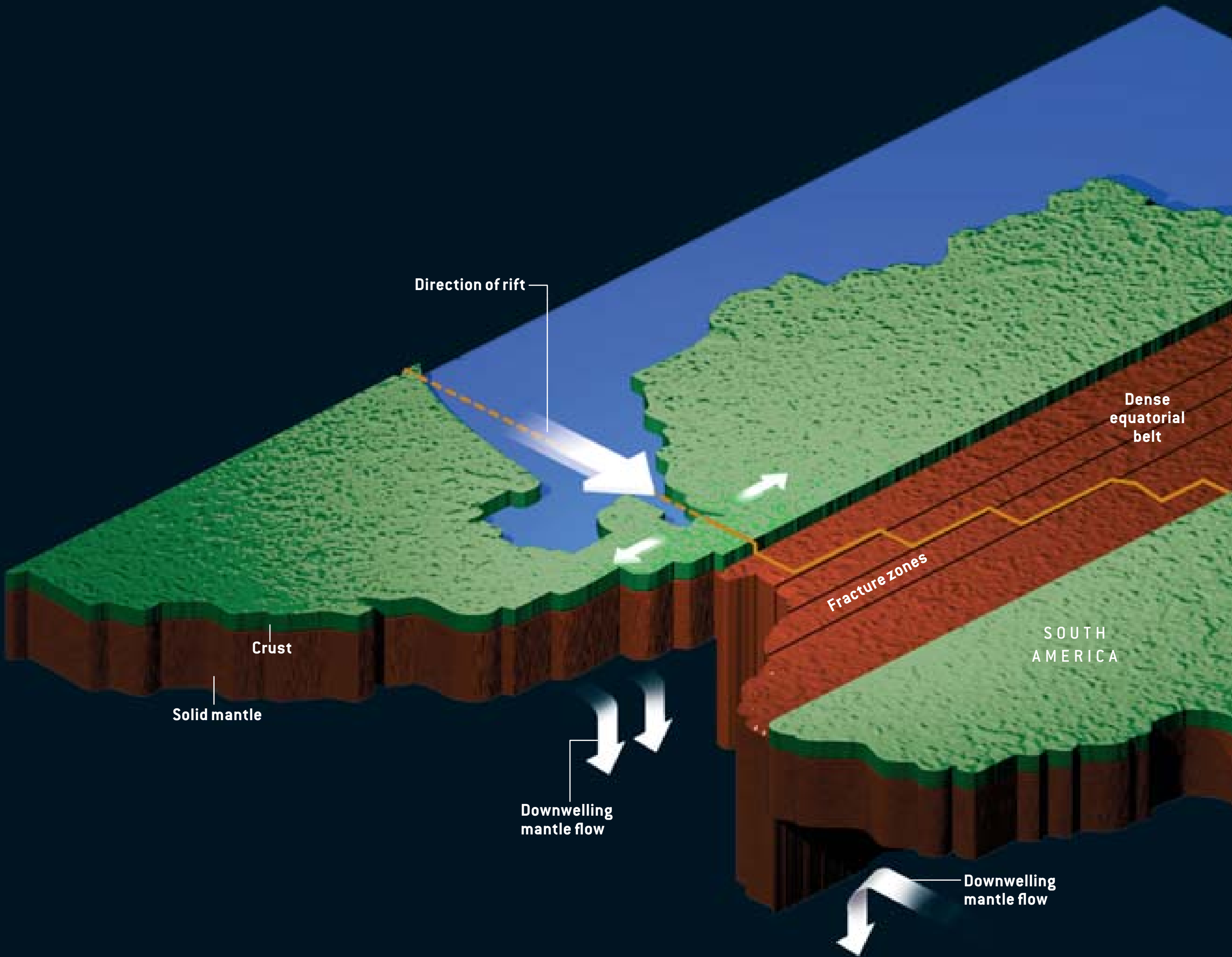
Dynamics of Cretaceous Vertical Motion of Australia and the Australian-Antarctic Discordance. Michael Gurnis, et al. in *Science*, Vol. 279, pages 1499–1504; March 6, 1998.

Dynamic Earth: Plates, Plumes and Mantle Convection. Geoffrey F. Davies. Cambridge University Press, 2000.

Constraining Mantle Density Structure Using Geological Evidence of Surface Uplift Rates: The Case of the African Superplume. Michael Gurnis et al. in *Geochemistry, Geophysics, Geosystems*, Vol. 1, Paper No. 1999GC000035; 2000. Available at **www.gps.caltech.edu/ffgurnis/Papers/Gurnis_et_al_Africa_1999.pdf**

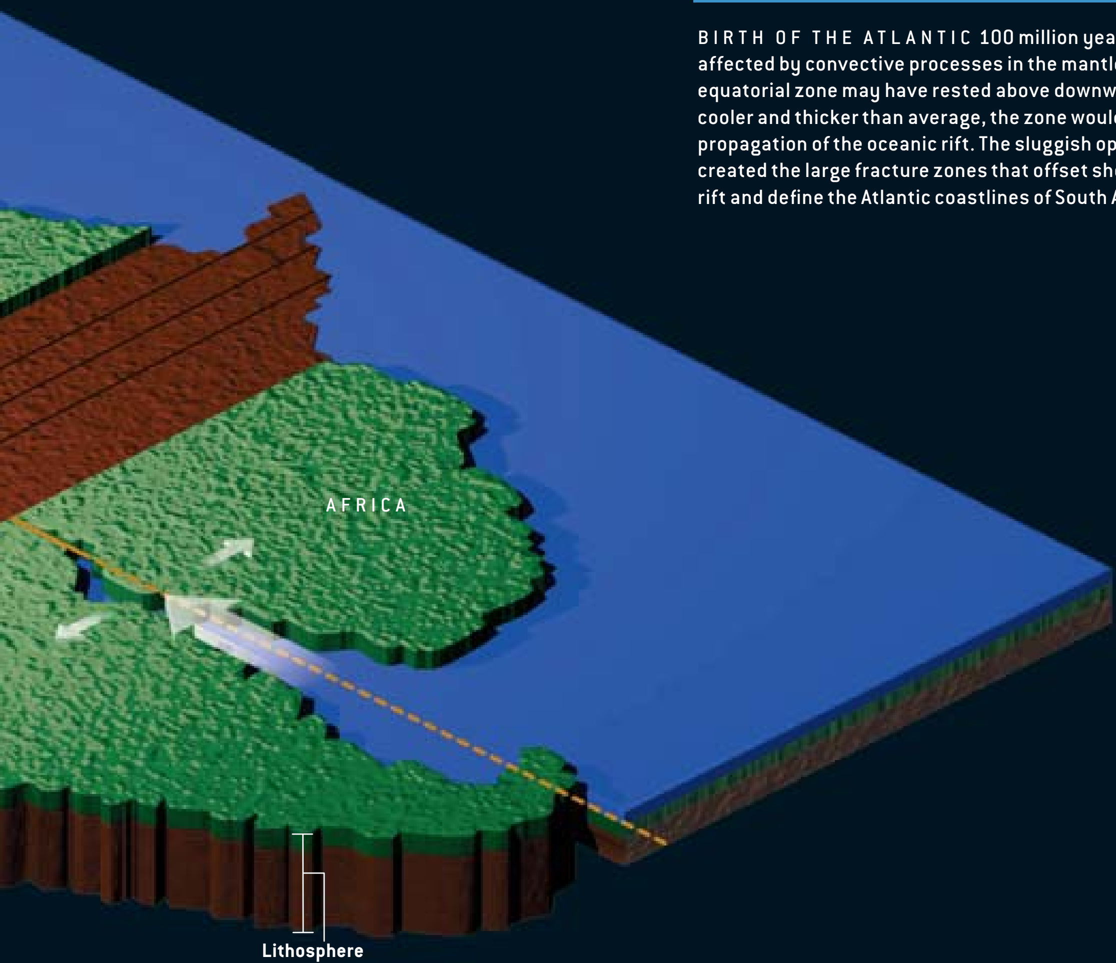
Computational Infrastructure for Geodynamics Web site: **www.geodynamics.org**

Gurnis's Computational Geodynamics Research Group Web site: **www.gps.caltech.edu/~gurnis/geodynamics.html**



Earth's Mantle below the Oceans

By Enrico Bonatti



BIRTH OF THE ATLANTIC 100 million years ago may have been affected by convective processes in the mantle. The lithosphere in the equatorial zone may have rested above downwelling mantle. Being cooler and thicker than average, the zone would have resisted the propagation of the oceanic rift. The sluggish opening would have created the large fracture zones that offset short segments of the rift and define the Atlantic coastlines of South America and Africa.

Samples collected from the ocean floor reveal how the mantle's convective forces shape Earth's surface, create its crust and perhaps even affect its rotation

Looking at a globe, one can easily imagine the continents and oceans as eternal, unchanging aspects of Earth's surface. Geophysicists now know that the appearance of permanence is an illusion caused by the brevity of the human life span. Over millions of years, blocks of Earth's rigid outer layer, the lithosphere, move about, diverging at mid-ocean ridges, sliding about along faults and colliding at the margins of some of the oceans. Those motions cause continental drift and determine the global distribution of earthquakes, volcanoes and mountain ranges.

Although the theory of plate tectonics is well established, the engine that drives the motion of the lithospheric plates continues to defy easy analysis because it is so utterly hidden from view. To confront that difficulty, I and other investigators have focused our research on the mid-ocean ridges. The ridges are major, striking locations where the ocean floor is ripping apart. Examination of the composition, topography and seismic structure of the regions along the mid-ocean ridges is yielding results that often run contrary to conventional expectations. More complicated and fascinating than anyone had anticipated, the chemical and thermal processes in the mantle below mid-ocean ridges dictate how new oceanic crust forms. Mantle activity may also cause islands to emerge in the middle of oceans and deep trenches to form at their edges. In fact, these processes may be so potent that they may even subtly affect the rotation of the planet.

The idea that Earth incorporates a dynamic interior may actually have its roots in the 17th century. In his 1644 treatise *Principles of Philosophy,* the great French philosopher René Descartes wrote that Earth had a central nucleus made of a hot primordial, sunlike fluid surrounded by a solid, opaque layer. Succeeding concentric layers of rock, metal, water and air made up the rest of the planet.

Geophysicists still subscribe to the notion of a layered Earth. In the current view, Earth possesses a solid inner core and a molten outer core. Both consist of iron-rich alloys and have temperatures reaching over 5,000 degrees Celsius and pressures well over a million times the pressure at the surface. Earth's composition changes abruptly about 2,900 kilometers below the surface, where the core gives way to a mantle much less dense than the core and made of solid magnesium-iron silicate minerals. Another significant discontinuity, located 670 kilometers below the surface, marks the boundary between the upper and lower mantle (the lattice structure of the mantle minerals changes across that boundary because of the different pressure). An additional major transition known as the Mohorovicic discontinuity, or Moho, separates the dense mantle from the lighter crust above it. The Moho lies 30 to 50 kilometers below the surface of the continents and less than 10 kilometers below the seafloor in the ocean basins. The lithosphere, which includes the crust and the upper part of the mantle, behaves like a mosaic of rigid plates lying above a hotter, more pliable lower part of the mantle called the asthenosphere.

THE AUTHOR

ENRICO BONATTI holds degrees in geology from the University of Pisa and the Scuola Normale Superiore in Pisa, Italy. After coming to the U.S. in 1959, he spent several years as a research scientist in marine geology at the University of California's Scripps Institution of Oceanography and as a professor at the University of Miami's Rosenstiel School of Marine Sciences. Since 1975 he has been with Columbia University's Lamont-Doherty Earth Observatory. Recently he has been teaching and researching in his native country. He has led or participated in expeditions in all the major oceans and in some remote but geologically intriguing lands, from the polar Ural region of Russia to the desert island of Zabargad in the Red Sea. Bonatti wishes to thank Daniele Brunelli, Anna Cipriani and Marco Ligi, who have collaborated with him in his research during the past decade.

MAKING RIDGES FROM MANTLE

THIS ORDERED, layered structure might seem to imply that Earth's interior is static. On the contrary, the deep Earth is quite dynamic. Thermal energy left over from the time of Earth's formation, augmented by energy released through the radioactive decay of elements such as potassium 40, uranium and thorium, churns the material within Earth. The heat travels across Earth's inner boundaries and sets into motion huge convection currents that carry hot regions upward and cold ones downward. These processes ultimately cause many of the broad geologic phenomena on the surface, including mountain building, volcanism, earthquakes and the motions of continents.

Heat travels across Earth's inner boundaries and sets into motion huge convection currents.

Among the regions offering the best access to Earth's interior are mid-ocean ridges. These ridges dissect all the major oceans, winding around the globe like the seams of a tennis ball, for a total of more than 60,000 kilometers. The Mid-Atlantic Ridge is a part of that global ridge system. A huge north-south scar in the ocean floor, it forms as the eastern and western parts of the Atlantic move apart at a speed of one to two centimeters per year. In addition to the frequent earthquakes that take place there, the summit of the Mid-Atlantic Ridge spews out hot magma during frequent volcanic eruptions. The magma cools and solidifies, thus forming new oceanic crust. The ridge is higher than the rest of the Atlantic basin floor. At progressively farther distances from the ridge, the seafloor deepens with respect to sea level, presumably because the lithospheric plates that move away from the ridge contract as they gradually cool with age.

The magma that rises at the Mid-Atlantic Ridge obviously originates in the upper mantle. Its composition differs considerably from that of the mantle, however. Magma that cools at ocean ridges forms a common kind of rock known as basalt. But researchers have found that seismic waves travel through the upper mantle at a rate of more than eight kilometers per second, far faster than they would pass through basalt.

One material that could possibly allow such a high velocity of sound is a type of dense, dark-green rock called peridotite. Peridotite consists mostly of three silicon-based minerals: olivine, a dense silicate containing magnesium and iron; orthopyroxene, a similar but less dense mineral; and clinopyroxene, which incorporates some aluminum and is more than 20 percent calcium. Peridotites also have small quantities of spinel, an oxide of chromium, aluminum, magnesium and iron.

How can basaltic magma be produced from a mantle made of peridotite? More than 30 years ago experimental

JACK HARRIS *Visual Logic (preceding pages)* petrologists such as Alfred E. Ringwood and David H. Green of the Australian National University exposed samples of peridotite to elevated temperatures (1,200 to 1,300 degrees C) and high pressures (more than 10,000 atmospheres). These values duplicate the temperature and pressure that exist in the suboceanic upper mantle roughly 100 kilometers below the seafloor. This research showed that gradual decompression of peridotite at those high temperatures melts up to 25 percent of the rock. The melt had a basaltic composition similar to that of mid-ocean ridge basalts.

These experiments support the view that hot, peridotitic mantle material rises under mid-ocean ridges from depths exceeding 100 kilometers below the seafloor. As the material moves upward, the mantle peridotite decompresses and partially melts. The melted part takes on the composition of a basaltic magma, rising rapidly toward the surface and separating from the peridotite that did not melt. Part of the melt erupts on the seafloor along the crest of the mid-ocean ridge, where it cools and solidifies and adds to the ridge crest. The remainder cools and solidifies slowly below the surface, giving rise to new oceanic crust. The thickness of the oceanic crust depends on the amount of melt that is extracted from the mantle.

The ridge crest's depth below sea level marks an equilibrium level determined by the temperature and initial composition of the upper mantle upwelling below the ridge. If the temperature and composition of the mantle were constant all along the ridge, the summit of the ridge would be at the same depth below sea level all along its length.

In the real world such consistency is unlikely. Small variations in mantle temperature or composition along the ridge would cause the summit to settle at varying elevations. Regions of suboceanic mantle where temperatures are higher have lower densities. In addition, a hotter mantle would melt more and produce a thicker basaltic crust. As a result, the ridge summits there will be higher.

The summit of the Mid-Atlantic Ridge shows just such variations in depth below sea level. For instance, along the ridge between about 35 and 45 degrees north latitude lies an area of abnormally high topography. Earth-orbiting satellites have detected in the same region an upward swell in the level of the geoid (the equilibrium level of Earth's surface, roughly equivalent to the average sea level).

Researchers generally attribute this swell to the influence of a so-called hot spot centered on the Azores island group. Hot spots are zones that have high topography and excess volcanism. They are generally interpreted as the surface expression of a "mantle plume"—that is, of a rising column of unusually hot mantle material. Most oceanic islands, including the Hawaiian Islands and Iceland, are thought to be the surface expressions of mantle plumes. The source of the heat is thought to lie in the boundary zones deep inside Earth, even as deep as the core-mantle boundary [see "The Core-Mantle Boundary," on page 36].

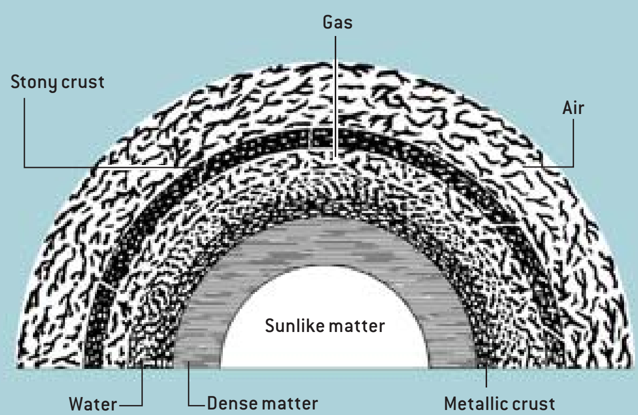
MINERALS OFFER EVIDENCE

MY COLLEAGUES AND I set out to test these ideas by exploring how the topography along the Mid-Atlantic Ridge relates to the temperature, structure and composition of the underlying mantle. One way to collect such information is to examine the velocities of seismic waves passing through the mantle under the ridge. Another approach involves searching for local variations in the chemistry of basalts that erupted along the axis of the ridge. Those variations can be used to infer the extent of melting and the physical nature of the mantle source from which they derived.

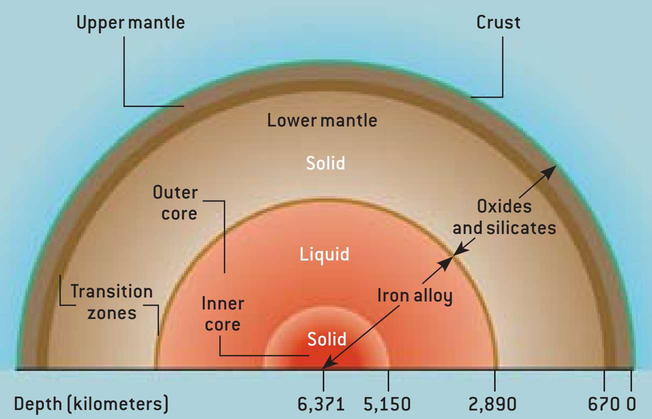
We followed a third approach by attempting to collect rock samples of mantle peridotite. Peridotite is left as a solid

EARTH'S LAYERED LOOK, THEN AND NOW

EARTH'S INTERIOR was imagined by French philosopher René Descartes in the 17th century (*left*). He viewed Earth as having a nucleus made of a hot, sunlike fluid covered by a dense, opaque solid. Succeeding layers consisted of metal, water, gas, stone and air.



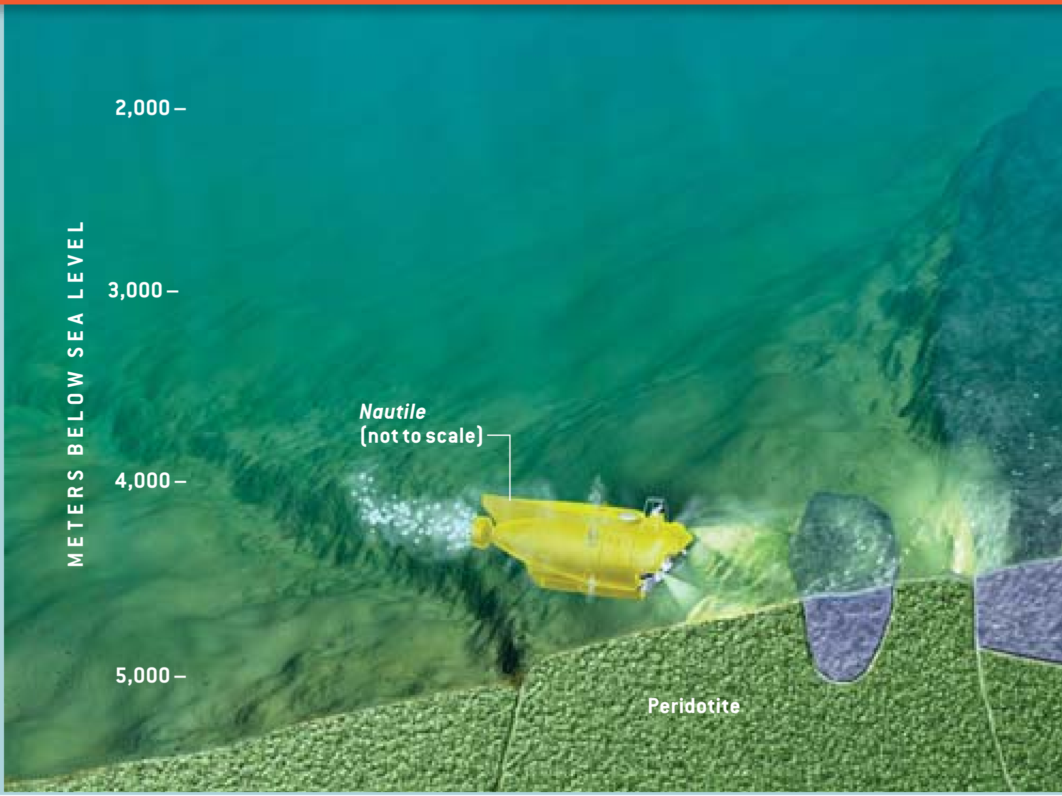
In the modern view (*right*), a solid inner core is cloaked by a molten outer core; both are made of iron alloy. The mantle is composed mostly of solid silicates and oxides of iron and magnesium.



GABOR KISS (*left*); DIMITRY SCHIDLOVSKY (*right*)

SURFING THE SEAFLOOR'S RIDGES AND VALLEYS

EXPLORATION OF THE SEAFLOOR by the *Nautile* occurred at the Vema Transform Fault, a valley that lies in the northern section of the Mid-Atlantic Ridge. Along the southern wall, mantle peridotites were found to outcrop in the lower part of the slope. Above them were gabbros, rocks created by the slow cooling of basaltic melt (the melted part of peridotite). The *Nautile* also discovered a dike complex, formed when basaltic melt cools and solidifies before reaching the seafloor. Above the dike complex lay pillow basalt, the form taken by basaltic melt that erupts on the seafloor and cools rapidly on contact with ocean water.



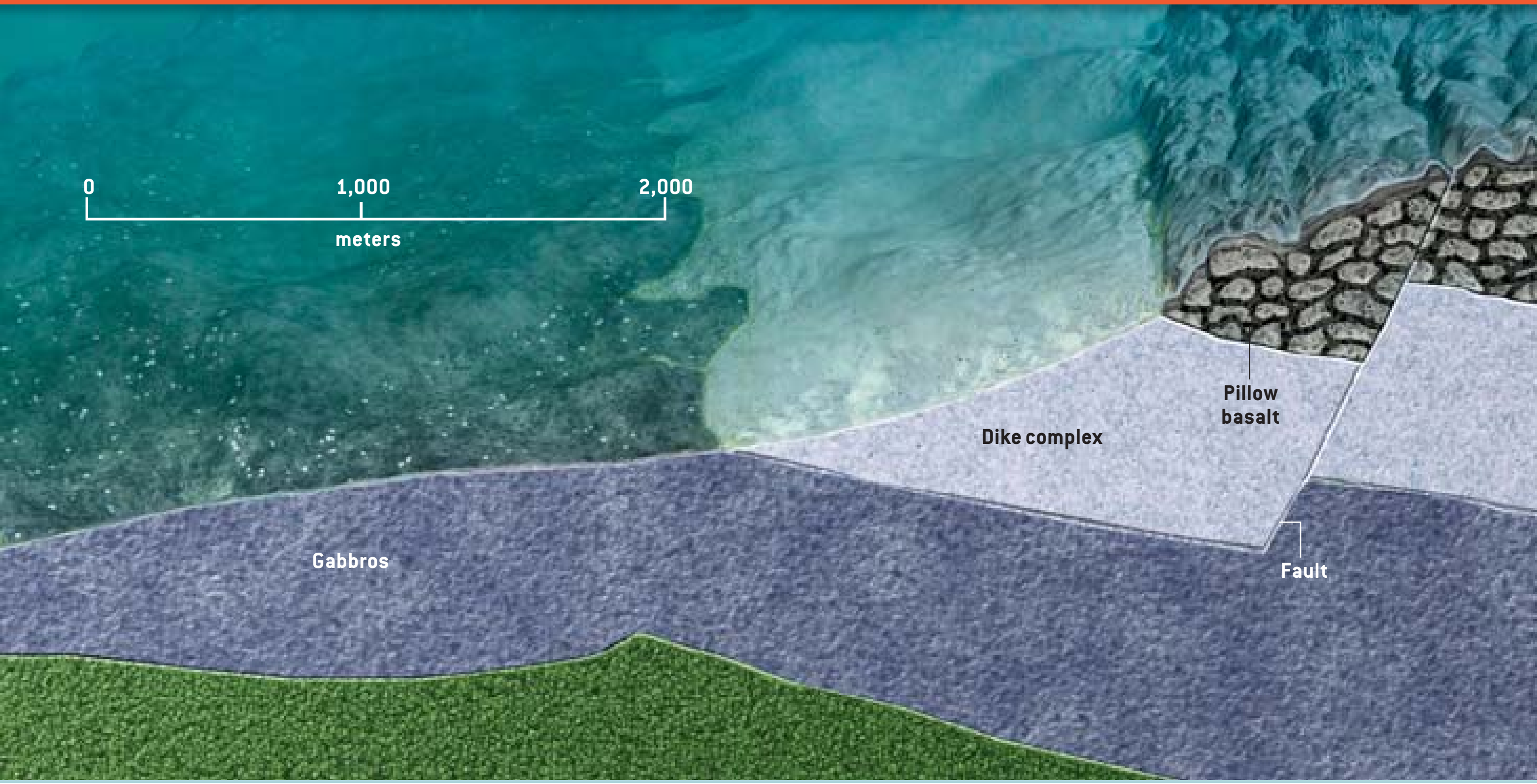
residue after the basaltic magma component melts out of the upper mantle rocks. Mantle rocks usually lie buried under several kilometers of ocean crust, but in some cases blocks of upper mantle peridotite are accessible. They are typically exposed where the axis of the mid-ocean ridge is faulted or where it is offset laterally by transform faults; these rocks can be sampled by drilling or dredging or retrieved directly through the use of a submersible.

To analyze the mantle minerals in the Atlantic peridotite samples, we used an electron microprobe. This instrument focuses a beam of electrons only a few microns in diameter onto a slice of rock. In response, the mineral emits x-rays of characteristic wavelengths. An analysis of the wavelengths and intensities of these x-rays allows a determination of the chemical composition of the mineral. Collaborating with Nobumichi Shimizu of the Woods Hole Oceanographic Institution and Luisa Ottolini of the Italian Research Council in Pavia, we also used a different instrument—an ion microprobe—to determine the concentration of trace elements such as titanium, zirconium and rare-earth elements. The ion probe focuses a beam of ions onto a sample, which dislodges other ions in the sample for measurement. The method enabled us to determine the concentrations of trace elements down to a few parts per million.

Such analyses reveal much about the conditions in the mantle where the sample rocks formed, because the temperatures and pressures there produce distinct compositions in the peridotites. Petrologists, including Green and A. Lynton Jaques of Geoscience Australia, have shown that partial melting modifies the relative abundances of the original minerals in the peridotite. Some minerals, such as clinopyroxene, melt more easily than do others and hence decrease in abundance during the melting. Moreover, the partial melting process changes the composition of the original minerals: certain elements in them, such as aluminum and iron, tend to follow the melt. Their concentration in the minerals decreases as melting proceeds. Other elements, such as magnesium and chromium, tend to stay behind, so that the solid residue becomes enriched with them. Thus, as a result of partial melting, olivine becomes more magnesium-rich and iron-poor; the ratio of chromium to aluminum in spinel increases; and so on.

The composition of these minerals, calibrated by laboratory experiments, allows us to estimate the degree of melting that mantle peridotites undergo during their ascent below the ridge. Our data showed that substantial regional variations exist in the composition of the mantle. For instance, the chromium-to-aluminum ratio of orthopyroxene and spinel is highest in peridotites sampled from a broad area between about 35 degrees and 45 degrees north latitude. The ratio suggests that the average degree of melting of the upper mantle lying below this region may reach as high as 15 percent. In most parts, about 10 to 12 percent of the mantle melts during the trip upward. This area of above-average melting corresponds to the Azores hot-spot region, lending credibility to the theory that hot spots result from unusually hot mantle plumes upwelling deep within Earth. Other find-

JACK HARRIS *Visual Logic*



ings support that idea, including work by Emily M. Klein, along with Charles H. Langmuir of Columbia University's Lamont-Doherty Earth Observatory, who independently examined the chemistry of basalts along the Mid-Atlantic Ridge.

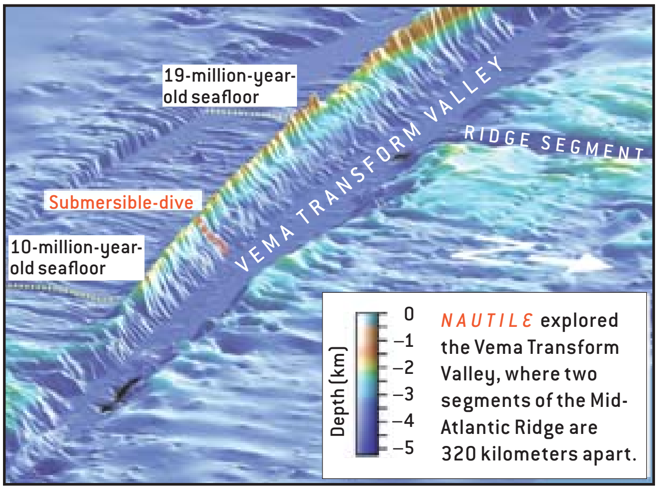
A hot spot would seem to be the cause of so much melting. In fact, assuming that temperature alone causes the melting in the Azores hot-spot region, we calculated that the hot-spot mantle would need to be more than 100 degrees C hotter than the mantle from elsewhere below the ridge.

Is there a way of testing the validity of this temperature estimate and its underlying assumptions? A number of geothermometers have been proposed. They are based on the observation that certain mineral pairs that coexist in equilibrium in the mantle undergo temperature-dependent chemical reactions. For instance, the orthopyroxene and clinopyroxene in a mantle peridotite react with each other until they reach an equilibrium composition that depends on temperature. Laboratory experiments have calibrated that relation. Thus, determining the composition of the coexisting mineral pair can indicate the temperature at which the members of the pair reached equilibrium.

I applied two geothermometers, one devised by Donald H. Lindsley of Stony Brook University and the other by Peter R. A. Wells of the University of Oxford, to the Mid-Atlantic Ridge peridotites. The results were surprising. They did not show higher temperatures in the hot-spot region; if anything, the region gives temperatures that are slightly lower.

A MANTLE SPIKED WITH WATER

WHY DID WE NOT FIND higher mantle temperatures for a region that displays high melting? One possibility is that the upper mantle there has a composition that causes it to melt more easily. Water could be the main factor. Experiments by Peter J. Wyllie of the California Institute of Technology and Ikuo Kushiro of the University of Tokyo and the Carnegie Institution of Washington, among others, demonstrated that trace amounts of water and other volatile elements in peridotite drastically decrease its melting temperature. Therefore, if such a "wet" mantle upwelled under a stretch of mid-ocean ridge, it would start melting more deeply in Earth than nor-



NAUTILE explored the Vema Transform Valley, where two segments of the Mid-Atlantic Ridge are 320 kilometers apart.

COURTESY OF ENRICO BONATTI

mal, "dry" mantle would. By the time the peridotite reached the surface, it would have undergone a degree of melting significantly greater than that of dry mantle under similar temperatures [*see box on opposite page*].

Is there any evidence that the upper mantle below the Azores hot-spot area is wetter than the mantle elsewhere below the Mid-Atlantic Ridge? Indeed there is. Jean-Guy E. Schilling and his co-workers at the University of Rhode Island reported that basalts from the segment of the hot spot situated between 35 and 45 degrees north latitude contain three to four times more water than do normal mid-ocean ridge basalts, as well as higher concentrations of several chemical elements (mostly light rare-earth elements). The anomalously high concentration of those elements means that the parent mantle in the hot-spot area harbors an enriched supply of these elements.

It seems, therefore, that the mantle below the Azores hot spot differs from the normal sub-Mid-Atlantic Ridge mantle not so much by being hotter as by having incorporated at some stage water and other fluids that changed its chemical composition and melting behavior. This chemical transformation of mantle peridotite by fluids is called metasomatism. It would explain why wet mantle near the surface would have experienced more melting than normal mantle would. It may also explain why the equilibrium temperatures estimated from peridotites at the Azores hot spot do not appear higher than average. Melting reactions consume heat, so that partial melting of upwelling mantle may actually have cooled the surrounding mantle. The higher the degree of melting, the greater the heat loss.

So the Azores hot spot may not be linked to a thermal plume originating from the deep mantle or the core-mantle boundary. Instead it may be a melting anomaly of relatively superficial origin in the mantle. These hot spots may not be truly hot and perhaps are best classified as "wet spots" because of the key role that fluids may play in their formation.

Where does the water that produces mantle metasomatism come from? One possible source of this water is the sinking slabs of old oceanic lithosphere in subduction zones at the margin of the oceans. This process recycles water into the mantle. Water could also be released in the upper mantle during degassing processes of the deep mantle. In addition, water molecules can be stored in the actual structure of mantle minerals.

Consider the mineral perovskite, a silicate of magnesium and iron that constitutes the main component of the lower mantle and is therefore the most abundant mineral on Earth. Perovskite can contain water in concentrations up to 1 percent. A lower-pressure form of perovskite, called wadsleyite, prevails in the zone of the mantle at a depth of between 660 and 450 kilometers and can contain water up to concentrations of about 1.5 percent. Totaling up all these water molecules, we can speculate that the total amount of water in Earth's mantle could be equivalent to that of several oceans. Much of this water is probably primordial, captured in Earth's mantle at the time of its formation over four billion years ago. The presence of water molecules dispersed in mantle minerals has important consequences. For example, it significantly lowers mantle viscosity, facilitating convective motions that cause the movement of lithospheric plates and the drifting of continents.

The water in Earth's mantle could equal the amount contained in several oceans.

UNEVEN MANTLE TILTS EARTH'S AXIS

OUR STUDIES OF MANTLE PERIDOTITES from the Mid-Atlantic Ridge suggest that some areas with cooler mantle temperatures may represent the return strokes of the convection cycle in the mantle—that is, the downwelling regions. To understand this notion, we must look south of the Azores region, to the equatorial zone where the Mid-Atlantic Ridge lies deeper than the ridge at higher latitudes. The mineral composition of peridotites recovered from the equatorial Atlantic indicates that they underwent little or no melting, which implies that the mantle temperature was exceptionally low. Schilling and Nadia Sushevskaya of the Vernadsky Institute of Geochemistry of the Russian Academy of Sciences reached similar conclusions after studying basalts from the equatorial Atlantic. In addition, Yu-Shen Zhang and Toshiro Tanimoto of Caltech found that the velocity of the seismic waves is faster in the upper mantle below the equatorial Mid-Atlantic Ridge than at higher latitudes. These observations imply a denser, colder upper mantle below the equatorial region of the Atlantic. The temperature of the upper mantle there may be up to 100 degrees C lower than the mantle temperatures elsewhere below the ridge.

A plausible explanation for the relatively cool and dense equatorial upper mantle is that it results from downwelling mantle currents. Hot mantle plumes upwelling in the northern and southern Atlantic mantle may flow toward the equator, giving up their heat to their cooler surroundings and then sink.

The equatorial position of the "cold" Atlantic mantle belt may not be arbitrary. It is possible that Earth's rotation and convection in the mantle are intimately connected phenomena. In the late 1800s George Darwin (the second son of Charles) pointed out that the distribution of large masses on the surface (such as continents) affects the position of Earth's axis of rotation. Several scientists since then have investigated how density inhomogeneities within the mantle cause true polar wander (that is, the shifting of Earth's axis of rotation relative to the mantle). The wander results from the natural tendency of a spinning object to minimize the energy spent for its rotation.

The redistribution of mass inside Earth may be recorded in the mantle. The late H. William Menard and LeRoy M. Dorman of the Scripps Institution of Oceanography suggested that the depth of mid-ocean ridges generally depends on latitude: ridges become deeper toward the equator and shallower toward the poles. Moreover, gravity measurements revealed that an excess of mass sits below the equatorial areas, at least in the Atlantic. These data suggest that abnormally cold and dense masses exist in the equatorial upper mantle.

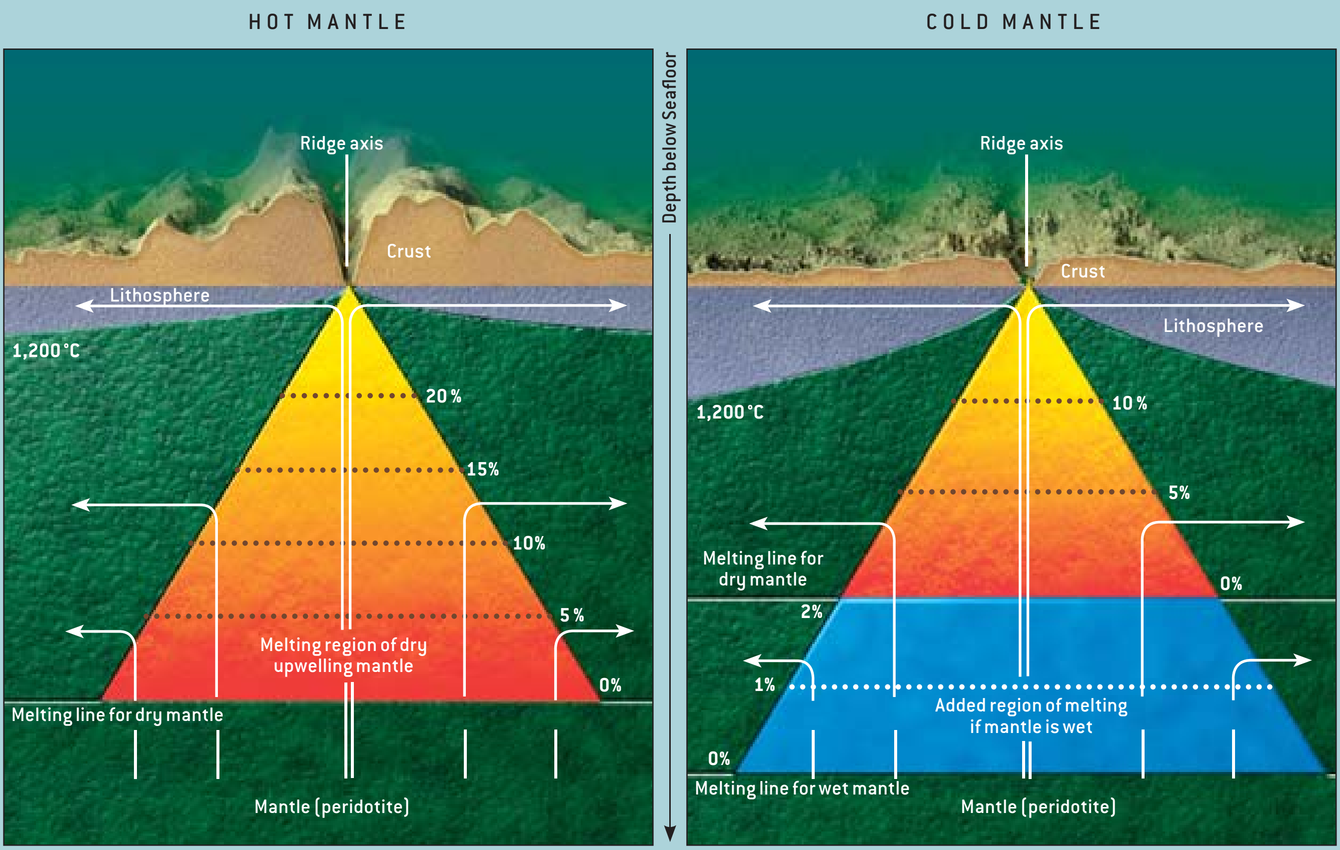
The sinking of cold, dense slabs into the mantle may influence true polar wander. Dense masses that find their way to the mantle, such as those that occur in subduction zones at the edge of some oceans, will affect the position of the rotation axis. The equator would tend to shift toward the dense masses. If high-density masses tend to concentrate near the equator, downwelling and cooler mantle spots are most likely to prevail in the equatorial upper mantle, explaining at least qualitatively the cold upper mantle belt and resulting lack of normal melting in the equatorial zone of the Atlantic.

DIVING FOR DEEP-SEA DATA

A COLDER-THAN-NORMAL equatorial mantle when the Atlantic first opened would imply a colder and thicker continental lithosphere along the equatorial belt. (The equator 100 million years ago crossed the future Atlantic coastlines of Africa and South America roughly along the same position as it does today.) The cold and thick equatorial lithosphere must have resisted the rift propagating from both the south and the north. The equatorial region may have behaved as a "locked zone" (in the sense used by French geologist Vincent E. Courtillot). As a result, the equatorial Atlantic opened sluggishly. This slow, difficult opening may have created the large equatorial fracture zones, visible today as east-west breaks that offset short segments of the mid-ocean ridge.

TEMPERATURE'S INFLUENCE ON MANTLE MELTING

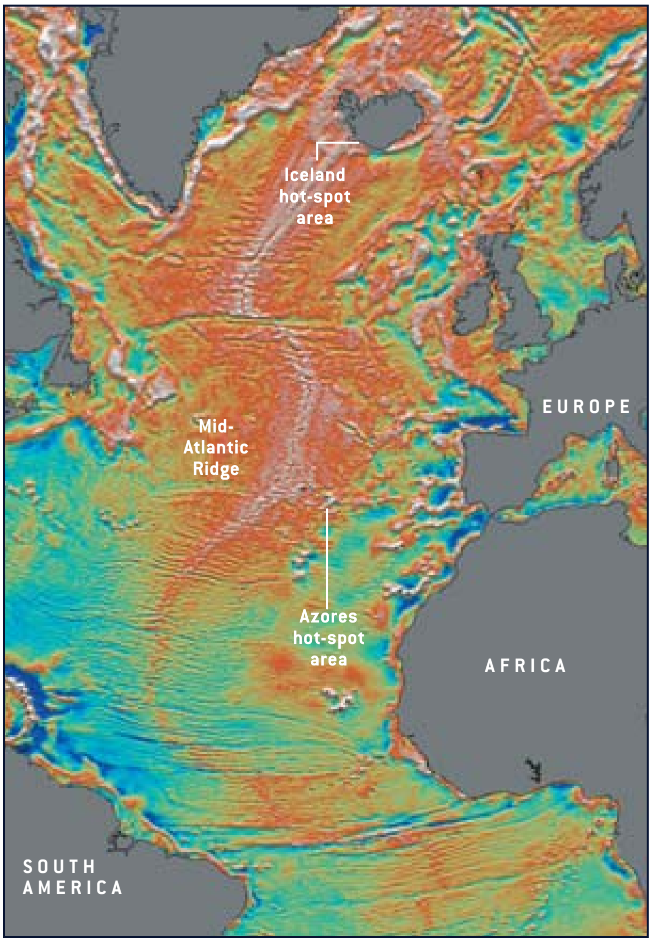
UPWELLING MANTLE melts to an extent that depends on whether the mantle is hot (*left*) or cold (*right*). The percentages indicate the amount of peridotite that melts. Melting proceeds until the peridotite stops rising and starts flowing horizontally. The hotter the mantle, the deeper the melting begins. As a result, more of the mantle melts, creating a thicker crust. Cold mantle melts less, unless it harbors fluids. In that case, it begins to melt much more deeply in Earth and may even melt more than hot mantle can. Wet mantle may explain why the Azores hot spot is rather cool.



DAVID FIERSTEIN

Now that we know that today's mantle upwelling below mid-ocean ridges is heterogeneous in terms of temperature and composition, the next question is: How do the properties of the mantle upwelling below a given segment of the ridge change over time? This information would enlighten us on an important issue, namely, how ocean basins evolve. But the research to obtain the necessary data would require sampling older oceanic lithosphere at various distances from the axis of the mid-ocean ridge. And unfortunately, the older lithosphere is normally buried deep below sediments.

We felt we might have a chance to reach old lithospheric material in the central Atlantic in the vicinity of the Vema Transform Fault. This fault offsets the crest of the Mid-Atlantic Ridge by 320 kilometers, cutting a deep valley through the oceanic crust. A long sliver of seafloor appears to have been uplifted on the southern side of the transform, and we hoped that this uplifted seafloor would expose a pristine section of lithosphere.



SATELLITE RADAR IMAGE of the North Atlantic reveals the topography of the seafloor. The radar measures variations in sea level, which correlate with the bumps and depressions of the seafloor. The Mid-Atlantic Ridge is clearly visible; it rises into broad swells associated with the Iceland and Azores hot spots. A large horizontal fracture interrupts the ridge between the hot spots. At the bottom of the image, long horizontal fractures traverse the Atlantic's "cold" equatorial seafloor between South America and Africa.

WILLIAM F. HAXBY *Lamont-Doherty Earth Observatory*

To test this hypothesis, in 1989 we organized an expedition in conjunction with Jean-Marie Auzende of the French oceanographic institution Ifremer. We planned to descend to the seafloor—more than five kilometers down—in the research submersible *Nautile*. Most of our colleagues viewed our task with skepticism: prevalent opinion held that the normal sequence of upper mantle and crust is completely disrupted near a transform fault.

Nevertheless, we pressed on. We began a series of dives that started at the base of the section and moved up the slope. Each dive lasted about 12 hours, about half of which was spent descending to the seafloor and returning to the surface. The cramped quarters of the *Nautile,* a sphere of titanium one meter and 80 centimeters in diameter, can accommodate two pilots and one scientist, who lies face down for the duration of the trip.

On our first dive we verified that the base of the section consists of mantle peridotite for a thickness of about one kilometer. On the second day we discovered a layer of gabbros—rocks that form below the seafloor when basaltic melts cool slowly—resting above the peridotite. According to widely accepted geophysical models, gabbros are the main component of the lower part of the oceanic crust. So in going upslope from mantle peridotites to crustal gabbros, we had crossed the Moho discontinuity.

The next day I took the *Nautile* on a dive that started from the level reached by the submersible the previous day. As I progressed along the slope, skimming the seafloor, a spectacular rock formation called a dike complex gradually revealed itself. Theory holds that dike complexes form where hot molten material, generated by partial melting of the mantle, squirts upward toward the seafloor through many narrow fissures in the crust. Never before had a dike complex been observed on the seafloor.

The dike complex, about one kilometer thick, was topped by a layer of pillow basalt, the form taken by basaltic magma when it cools and solidifies rapidly on eruption to the seafloor. During the next several days, we explored a different section and confirmed our previous findings. We were quite excited because no one had ever before observed a complete and relatively undisturbed section of oceanic upper mantle and crust. We immediately documented our discovery in a short paper that we mailed to *Nature* as soon as we docked a few weeks later.

Encouraged by the results of the *Nautile* dives, we conducted two other expeditions and established that the Vema lithospheric section is exposed on the seafloor for more than 300 kilometers. After mapping the magnetic anomalies produced by the seafloor, we could estimate the velocity with which the lithosphere moves away from the ridge axis. We thus established that the Vema section exposes lithosphere created gradually at the axis of the Mid-Atlantic Ridge during a time interval of more than 20 million years—a unique opportunity to study how the creation of lithosphere varies through time!

During the dives, we had used the *Nautile*'s mechanical arm to grab a number of samples of mantle peridotite. We later sampled by dredging mantle peridotites at close intervals along the base of the section in lithosphere of increasing age. From the mineral composition of these rocks we estimated the variations in the degree of melting they had undergone over time during their ascent below the Mid-Atlantic Ridge. At the same time, we could estimate how crustal thickness varied through time, thanks to gravimetric data obtained from both ship and satellite measurements of the gravity field produced by rocks below the seafloor. Crustal thickness depends on the quantity of melt generated by mantle ascending below the ridge.

The results were quite unexpected. The degree of melting of the mantle and the crustal thickness both appear to have increased steadily from 20 million years ago to today. Small oscillations are superimposed on this general trend. The simplest interpretation of these results: the Mid-Atlantic Ridge is becoming steadily "hotter" over time.

Surprisingly, the increase of the temperature of the upwelling mantle is accompanied by a decrease in the spreading rate of the lithospheric plate generated at the ridge axis. This result contrasts with the concept of "passive" upwelling of the mantle in response to the diverging motion of the lithospheric plates—a concept that would require proportionality between spreading rate and degree of melting of the ascending mantle.

We were also able to estimate the velocity of the solid mantle that rises below the ridge, crucial information for refining our models on the formation of the oceanic crust. The speed of the rising mantle depends on its temperature and composition (both affect density and viscosity) and on the diameter of the rising column and is related to the velocity of the lithospheric spreading that diverges from the ridge axis.

How can we estimate the speed of the rising solid mantle? The rising mantle generates melt within a depth interval that can be estimated from experiments and theoretical considerations. The melt fraction rises rapidly, cooling and solidifying as basalt in the crust, while its parent mantle continues to ascend slowly.

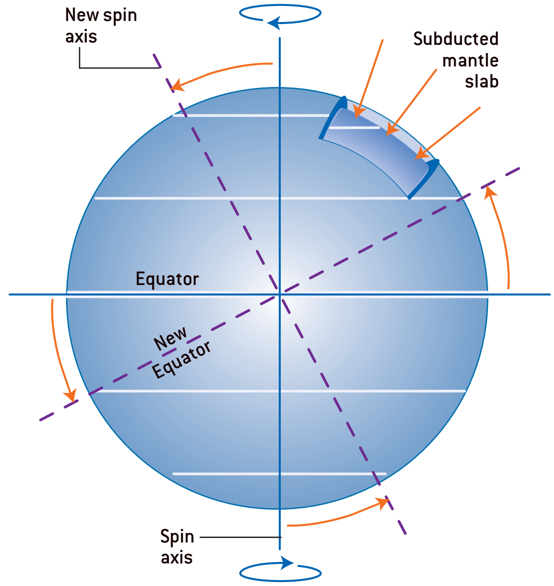
When the "parent" mantle peridotite reaches the lithosphere and starts moving horizontally with the plate away from the ridge, the basalt it generated has moved farther away from the ridge. The horizontal distance between the parcel of basaltic crust and its parent mantle, translated as time, would allow us to estimate the velocity of the rising solid mantle. After correlating the temporal variations of the degree of mantle melting with the variations of crustal thickness along the Vema lithospheric section, we estimated the solid mantle rose at an average velocity of about 25 millimeters per year.

To refine this estimate, we need to go back and take additional samples of peridotite from the exposed lithospheric section so that we can achieve a higher resolution in the curve describing temporal variations of degree of melting of the mantle.

Why is the Mid-Atlantic Ridge north of the equator becoming gradually hotter? We can only speculate. Perhaps a wave of plume-derived hot mantle has been flowing southward toward the equator since a few tens of million years ago. We have hints that major oscillations in the intensity of mid-ocean ridge activity occurred in the distant past.

For example, studies by Roger Larson of the University of Rhode Island suggest that a mantle "superplume" roughly 100 million years ago caused swelling of mid-ocean ridges, faster seafloor spreading, rising sea levels, and warming of the climate as a result of larger quantities of carbon dioxide, methane and other greenhouse gases released from the mantle [see "The Mid-Cretaceous Superplume Episode," on page 22].

Much remains to be done before geologists develop a complete picture of mantle dynamics and its influence on surface geology. Debate persists as to the origins of mantle convection and whether it extends into the lower mantle. Indeed, symposia that include theoreticians, geophysicists, geochemists and petrologists invariably yield heated discussions and much dissent. On one point there is unanimity: Earth's mantle is very much alive and is an exciting region to study. SA



SHIFTING OF EARTH'S AXIS can be influenced by the sinking of cold, dense slabs of mantle. Such sinking occurs in subduction zones, such as those surrounding the Pacific Ocean. Earth's axis of rotation would tend to shift so that the equator would move closer to the dense slabs.

DIMITRY SCHIDLOVSKY

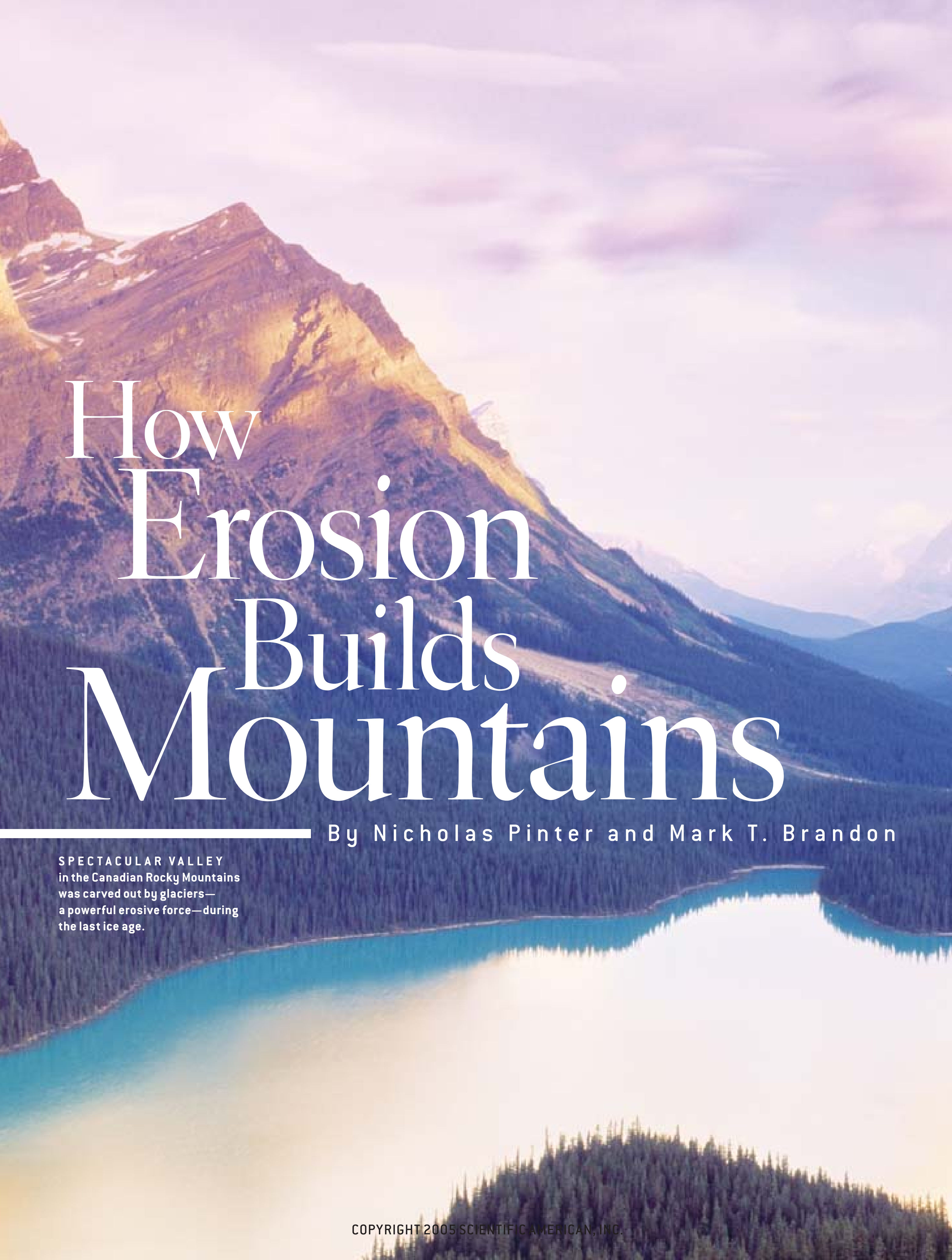
MORE TO EXPLORE

Ridges, Hotspots and Their Interaction as Observed in Seismic Velocity Maps. Y. S. Zhang and T. Tanimoto in *Nature*, Vol. 355, pages 45–49; January 2, 1992.

A Cold Suboceanic Mantle Belt at the Earth's Equator. E. Bonatti, M. Seyler and N. Sushevskaya in *Science*, Vol. 261, pages 315–320; July 16, 1993.

Mantle Thermal Pulses below the Mid-Atlantic Ridge and Temporal Variations in the Formation of Oceanic Lithosphere. E. Bonatti, M. Ligi, D. Brunelli, A. Cipriani, P. Fabretti, V. Ferrante, L. Gasperini and L. Ottolini in *Nature*, Vol. 423, pages 499–505; May 29, 2003.

Water Rich Basalts at Mid-Ocean Ridge Cold Spots. M. Ligi, E. Bonatti, A. Cipriani and L. Ottolini in *Nature*, Vol. 434, pages 66–69; March 3, 2005.



How Erosion Builds Mountains

By Nicholas Pinter and Mark T. Brandon

SPECTACULAR VALLEY in the Canadian Rocky Mountains was carved out by glaciers—a powerful erosive force—during the last ice age.

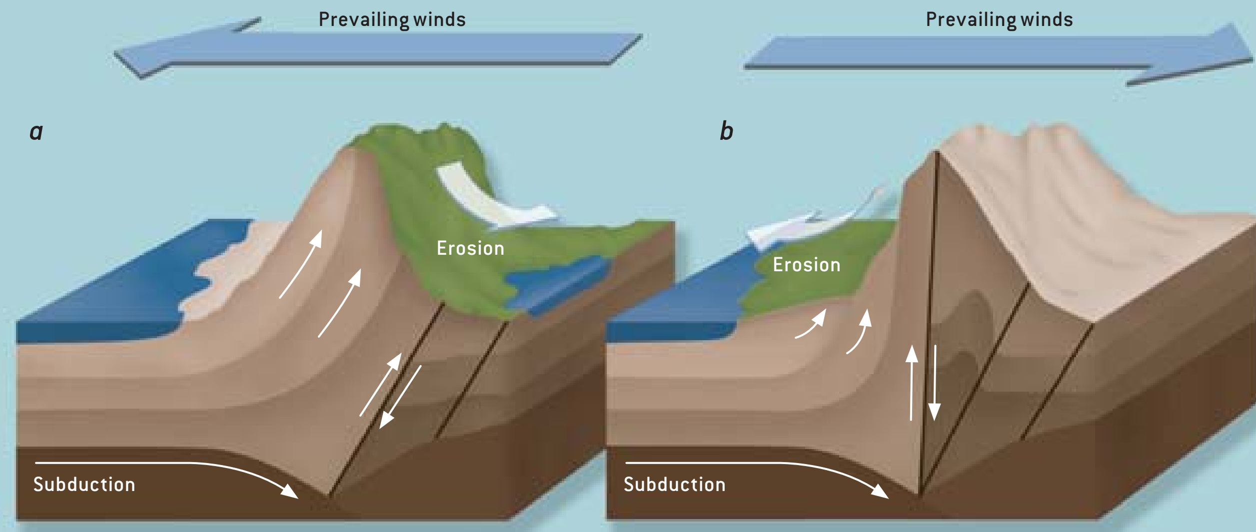
RICHARD SISK *Panoramic Images*

An understanding of how tectonic, erosional and climatic forces interact to shape mountains permits clearer insights into Earth's history

Mountains have evoked awe and inspired artists and adventurers throughout human existence. Recent research has led to important new insights into how these most magnificent of Earth's formations came to be. Mountains are created and shaped, it appears, not only by the movements of the vast tectonic plates that make up Earth's exterior but also by climate and erosion. In particular, the interactions between tectonic, climatic and erosional processes exert strong control over the shape and maximum height of mountains as well as the amount of time necessary to build—or destroy—a mountain range. Paradoxically, the shaping of mountains seems to depend as much on the destructive forces of erosion as on the constructive power of tectonics. In fact, after 100 years of viewing erosion as the weak sibling of tectonics, many geologists now believe erosion actually may be the strong one in the family. In the words of one research group, "Savor the irony should mountains owe their [muscles] to the drumbeat of tiny raindrops."

Because of the importance of mountain building in the evolution of Earth, these findings have significant implications for earth science. To a geologist, Earth's plains, canyons and, especially, mountains reveal the outline of the planet's development over hundreds of millions of years. In this sprawling history, mountains indicate where events in or just below Earth's crust, such as the collisions of the tectonic plates, have thrust this surface layer skyward. Thus, mountains are the most visible manifestation of the powerful tectonic forces at work and the vast time spans over which those forces have operated.

TURNING WIND INTO RAIN



OROGRAPHY is the phenomenon in which mountains lift the air currents flowing over them, increasing precipitation over the crest and windward slopes of the range. In a mountain range near an ocean, for example, when the prevailing winds blow offshore, opposite to the direction of subduction (*a*), erosion is concentrated on the inland side of the range, exposing the deepest, most deformed rocks in that area.

When the wind is in the same direction as subduction (*b*), erosion denudes the coastal side of the range, literally pulling buried rocks toward the surface. In this case, the inland side of the mountain range lies in an arid "rain shadow," such as the desert east of the Sierra Nevada (*photograph*).

The effort to understand mountain building has a long history. One of the first comprehensive models of how mountains evolve over time was the Geographic Cycle, published in 1899. This model proposed a hypothetical life cycle for mountain ranges, from a violent birth caused by a brief but powerful spasm of tectonic uplift to a gradual slide into "old age" caused by slow but persistent erosion. The beauty and logic of the Geographic Cycle persuaded nearly a century of geologists to overlook its overwhelming limitations.

In the 1960s the plate tectonics revolution explained how mountain building is driven by the horizontal movements of vast blocks of the lithosphere—the relatively cool and brittle part of Earth's exterior. According to this broad framework, internal heat energy shapes the planet's surface by compressing, heating and breaking the lithosphere, which varies in thickness from 100 kilometers or less below the oceans to 200 kilometers or more below the continents. The lithosphere is not a solid shell but is subdivided into dozens of plates. Driven by heat from below, these plates move with respect to one another, accounting for most of our world's familiar surface features and phenomena, such as earthquakes, ocean basins and mountains.

Earth scientists have by no means discarded plate tectonics as a force in mountain building. Over the past few decades, however, they have come to the conclusion that mountains are best described not as the result of tectonics alone but rather as the products of a system that encompasses erosional and climatic processes in addition to tectonic ones and that has many complex linkages and feedbacks among those three components.

THE ROLE OF TECTONICS

PLATE TECTONICS still provides the basic framework that accounts for the distribution of mountains across Earth's surface. Mountain building is still explained as the addition of mass, heat or some combination of the two to an area of Earth's crust (the crust is the upper part of the lithosphere). Thicker or hotter crust rises upward, forming mountains, because the crust is essentially floating on the mantle under it, and crust that is either thicker or hotter (less dense) floats higher. Plate tectonics contributes to the thickening of the crust by either lateral convergence between adjacent plates or through the upward flow of heat and magma (molten rock).

Convergence of tectonic plates generally occurs in one of two ways. One plate may slide down, or subduct, below the other, into the mantle. At a subduction zone boundary, the

upper plate is thickened as a result of the compression and from magma being added by the melting of the descending plate. Many mountains, including almost all the ranges that surround the Pacific Ocean in a geologically active area known as the ring of fire, formed by subduction. With continental collision, on the other hand, neither plate subducts into the mantle, and therefore all the mass added as a result of the collision contributes to the building of mountains. Such collisions have created some spectacular topography, such as the Tibetan Plateau and the Himalayas, the mountain range that includes the world's 10 highest peaks.

The flow of magma and heat to Earth's crust—during volcanic activity, for example—can also drive mountain building. Earth's longest mountain chains—the mid-ocean ridges—are the result of magma welling up as adjacent plates move apart, forming new crust under the ocean. These ridges run through the Atlantic, eastern Pacific and Indian oceans like the seam on a baseball; the Mid-Atlantic Ridge alone is more than 15,000 kilometers long, rising as much as 4,000 meters above the surrounding abyssal plains of the ocean floor. On land, heat associated with the flow of magma can also help uplift large areas by making the lithosphere less dense and more buoyant on the underlying mantle.

CLIMATE AND EROSION

THE EMERGING, SYSTEM-ORIENTED VIEW of mountain building adds to these tectonic phenomena the often closely intertwined effects of erosion and climate. Erosion includes the disaggregation of bedrock, the stripping away of sediment from slopes and the transport of the sediment by rivers. The mix of erosional agents active on a particular landscape—gravity, water, wind and glacial ice—depends on the local climate, the steepness of the topography and the types of rock at or near the surface.

Climate is inextricably linked with erosion because it affects the average rate of material loss across a landscape. In general, wetter conditions favor faster rates of erosion; however, more moisture also promotes the growth of vegetation, which helps to "armor" the surface. Mountains in polar latitudes are the least vulnerable to erosion, partly because of the aridity of cold climates and partly because continental ice sheets such as those on Greenland and Antarctica commonly are frozen to the underlying rock and cause little erosion. In contrast, mountain glaciers such as those of the European Alps and the Sierra Nevada in California aggressively attack the subsurface rock, so that this type of glacier may be Earth's most potent erosional agent.

There are many other links among erosion, climate and topography. For example, mountains lift the winds that flow over them, causing increased precipitation on the range's windward slopes, intensifying erosion as a result. Known as orography, this effect is also responsible for the "rain shadow" that creates deserts on the leeward sides of many mountain ranges [*see photograph on opposite page*]. Elevation can also affect erosion, because average temperature decreases with altitude, so that higher peaks are less likely to be protected by vegetation and more likely to be eroded by glaciers. In temperate regions the rate of erosion is proportional to the average steepness of the topography, apparently because gravity- and water-driven processes are more effective on steeper slopes. Taken together, all these facts suggest that mountains evolve their own climates as they grow—becoming typically wetter, colder and characterized by more intense erosion.

The links described above demonstrate that mountain ranges are best viewed as a system. To understand the behavior of any such system, it is necessary to identify both its components and the interactions among those components. Because these interactions are so important, simple system inputs can lead to surprisingly complex outputs. Such complexities include feedback—stabilizing or destabilizing links between component processes. In the simple example we have outlined, the system is forced by tectonic collision, which adds mass to the mountain belt, and the response is an increase in the average height of the mountain range. As the mountains grow taller, erosion increases, reducing the growth rate. This example illustrates negative feedback, in which continued positive forcing of a system leads to a progressively reduced response. In contrast, positive feedback has the opposite effect, accelerating any change in a system. The creation of a rain shadow is an example of positive feedback; erosion is inhibited, allowing a mountain range to continue its rapid growth. The rain shadow north of the Himalayas

MEGAN O'DEA; SOURCE: SEAN WILLETT, CHRISTOPHER BEAUMONT AND PHILIPPE FULLSACK IN *GEOLOGY*, VOL. 21, NO. 4; APRIL 1993 [*illustration*]; GERALD FRENCH *Panoramic Images* [*photograph*]

has contributed to the formation of the high-standing Tibetan Plateau [*see box on pages 80 and 81*].

The concept of feedback is at the heart of the new understanding of how mountains are built—and even how mountain building affects the Earth system as a whole. Numerous different types of feedback have been recognized or postulated. Among the most unexpected insights that have accrued from these discoveries is the realization that several important feedbacks enable surface processes, such as climate and erosion, to influence profoundly tectonic processes deep below the surface (and vice versa).

ISOSTASY IS KEY

ONE IMPORTANT FEEDBACK occurs through the phenomenon known as isostasy, which refers to the buoyancy of Earth's crust as it floats on the denser, fluidlike mantle below it. A mountain range, like any physical structure, must be supported, and it turns out that this support comes mainly from the strength of the crust and from isostasy. Under the soaring peaks of every mountain range is a buoyant "root" of crust that penetrates into the mantle. Icebergs offer a useful analogy: because ice is about 90 percent as dense as water, a given mass of ice above the water is supported by nine times that mass underneath the waterline. Continental crust is about 80 to 85 percent as dense as the mantle beneath, enabling crustal roots tens of kilometers deep to support mountains several kilometers high.

Isostasy is the key mechanism that links a mountain's tectonic, or internal, evolution to its geomorphic, or external, development. When erosion at the surface removes mass, isostasy responds by lifting the entire mountain range up to replace about 80 percent of the mass removed. This uplift explains a number of phenomena that were puzzling before researchers fully appreciated the role of feedback in mountain building.

For example, high-precision surveys along the eastern margin of the U.S. have revealed that the land is rising at rates of a few millimeters to a few centimeters a century. This was puzzling because the Appalachian Mountains lie in the interior of the North American plate, where there is no convergent plate boundary to account for the uplift. Some geologists suggested that the survey results must therefore have been in error. Given our new understanding, however, some or all of the measured uplift may be the isostatic response to erosion, especially in the high-relief areas of the Appalachians. Erosion that is concentrated at the bottom of river valleys may be especially significant because it can lift mountain peaks to elevations *higher* than the elevations before erosion started. This is possible because the removal of mass is localized (in the valleys), but the isostatic response lifts the entire mountain block, including both valleys and peaks.

Although isostasy can prop them up for many millions of years, landscapes without tectonic uplift do eventually succumb to erosion. Several studies have suggested that large areas of Australia are good examples of very old, decaying landscapes. These areas, which have not experienced tectonic uplift for hundreds of millions of years, are at most a few hundred meters above sea level. Their rates of surface uplift seem to be consistent with only isostatic response to erosion. In such tectonically active mountains as the Himalayas and the European Alps, measured uplift reflects a combination of tectonic driving forces and erosionally driven isostatic uplift. Given the rates at which mountains grow and then decay, we can infer that dozens of major mountain ranges have come and gone on Earth throughout its history.

UNUSUAL TECTONIC TIMES?

THE CONSTRUCTION OF MOUNTAINS, including ancient mountains that were built and eroded away in the distant past, can leave a variety of marks in the geologic record, such as those from lava flows, intrusion of magma, the exposure of once deeply buried rocks, as well as copious sediment deposited in lowland basins and the fossils of plants known to thrive only at high altitudes. By studying such indicators from many different periods, geologists can make inferences about the extent of mountain building on

HIMALAYAS and Tibetan Plateau are visible in this satellite image as the mostly white areas north and east of India—towering manifestations of an ongoing collision that started 50 million years ago, when the Indian tectonic plate, moving north, began plowing into the Eurasian plate. The collision's most visible result is the high, flat topography of Tibet. In contrast, the Himalayas—the high snowcapped range along the southern margin of the plateau—comprise just a small fraction of the area created by this collision.

TOM VAN SANT *Geosphere Project*

THE AUTHORS

NICHOLAS PINTER and *MARK T. BRANDON* began their collaboration at Yale University in the emerging field of active tectonics, which emphasizes the interactions between tectonic deformation and Earth's topography. Pinter carried out postdoctoral research there and is now professor at Southern Illinois University at Carbondale. His research includes a focus on the topographic expression of tectonic processes and has involved work in California, South America and the peri-Adriatic region. Brandon is professor of structural geology and tectonics at Yale. His research is focused on understanding the interrelation between tectonic uplift and erosion at subduction zones and collisional mountain ranges. Areas he is studying include the Apennines, the Alps, the southern Andes, Crete and the Coast Ranges of the western U.S.

Earth at different times, thereby gaining insights into the planet's development.

Various geologists have looked at the relative abundance of sediment, magmatic activity and other potential indicators of mountain building and concluded that the past 40 million years represents an anomalous surge of tectonic activity and mountain building. This same geologic period, however, also saw a major climate shift on Earth, a global cooling that transformed Greenland and Antarctica from temperate, vegetated lands to permanent ice sheets and that culminated in the glaciers that covered North America and Europe during the past two million years. Given this evidence, two opposing theories have been proposed to explain mountain building and climate over the past 40 million years: either the surge of mountain building caused the global climate shift, or the climate shift caused the surge of mountain building.

The first of these two theories asserts that long-term cooling was caused by a surge in mountain building around the globe. For example, glaciers tend to be self-perpetuating: once established, they increase the reflectivity, or albedo, of the surface, thus lowering temperatures and allowing more ice to form. Widespread uplift of large mountain masses in the past 40 million years could have increased the area of Earth covered by mountain glaciers, which would have increased the albedo of the planet. Atmospheric carbon dioxide may have been another important feedback agent. One interpretation states that mountain building can alter the global distribution of rain and snowfall, increasing the pace at which rock is broken down by dissolution and chemical reactions. According to this hypothesis, accelerated chemical weathering removed carbon dioxide from the atmosphere, reducing the greenhouse effect and thereby leading to a cooler global climate.

The second theory about mountain building and climate contends that climate change was really the more powerful of those two forces during the past 40 million years. This theory proposes that climate change has actually produced many of the profound geologic changes that are usually attributed to accelerated mountain growth. Global cooling may have been driven by continental drift, which changed the distribution of land and ocean area with respect to latitude as well as the pattern of ocean currents, which are major mechanisms by which Earth equilibrates the heat imbalance between the equator and the poles [see "Chaotic Climate," by Wallace S. Broecker; SCIENTIFIC AMERICAN, November 1995]. How could these climate changes mimic mountain building? Through isostatic uplift. According to this interpretation, global cooling intensified erosion in many mountain ranges. Stepped-up erosion, particularly in the bottom of river and glacial valleys, resulted in increased uplift of mountain summits as isostasy compensated for the removal of mountain mass by erosion.

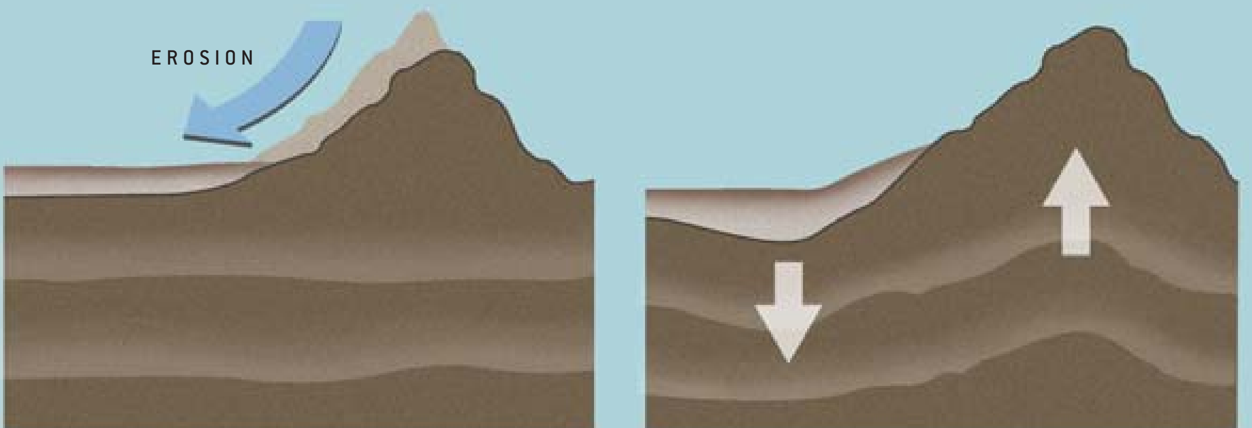
The cause-and-effect ambiguity between global climate and mountain building has been billed as a geologic paradox to rival the "chicken and egg" question, but such circularity is common in feedback-rich systems. Geologists may not currently know what initiated the changes in climate and topography that have occurred in the past 40 million years, but they now understand that the many kinds of feedback in this system are capable of amplifying any change and that tectonics, climate and erosion must have acted together in creating the geologic evidence that we find today.

EROSION'S PULL

RECOGNITION OF THE MANY TYPES of feedback in the mountain-building system reveals that erosion not only participates in shaping mountains but also guides tectonic processes deep within the crust. The ultimate limiting force to mountain growth is gravity. Thus, erosion, by reducing the weight of the mountain range, actually accelerates tectonic processes beneath the mountains. For this reason, erosional processes can be viewed as "sucking" crust into mountain

WHEN MOUNTAINS FLOAT ON MANTLE

ISOSTATIC UPLIFT occurs as a result of the buoyancy of a mountain on the more dense, fluidlike mantle (*not shown*) on which it "floats." Erosion causes the crust to rise up, whereas deposition of the resulting sediment in the basin area weighs the crust downward.



MEGAN O'DEA

THE HIMALAYAS AND THE APPALACHIANS

TWO OF EARTH'S grandest mountain ranges are the Himalayas and the Appalachians. Both were built by continental collisions, but the two ranges are about as different as mountains can be. Their comparison illustrates well the key principles of the new, system-oriented view of mountain building.

Stretching 2,500 kilometers across northern India and southern Tibet, the Himalayas are the king of Earth's mountain ranges. In this range stand many of the world's highest peaks, including Mount Everest, the tallest at 8,848 meters. Together with the Tibetan Plateau, to the north of the range in southwest China, the Himalayas contain the globe's greatest total mountain mass. It has even been suggested that this mountain belt is the largest high-elevation mass that Earth has seen in the past billion years. In spite of this range, the landscape atop the Tibetan Plateau is surprisingly flat. The plateau is Earth's largest expanse of land above 5,000 meters—a region approximately half the area of the continental U.S., most of it at least 600 meters higher than Mount Whitney, the highest single point in the continental U.S.

All this dramatic and varied topography developed during the past 50 million years, as a result of the collision between the Indian and the Eurasian tectonic plates. The collision began to squeeze both India and Tibet, activating a series of crustal-scale contractional faults that thrust part of the Indian continent underneath southern Asia. The northward velocity of India before the collision was 15 to 20 centimeters a year, and the velocity afterward was about five. Such deceleration of an entire continent is less surprising than the fact that India has continued to plow into and through southern Asia at about five centimeters a year for the past 40 million to 50 million years. India has advanced 2,000 kilometers into the Eurasian plate, give or take 800 kilometers, roughly doubling the thickness of the crust, uplifting the Himalayas and the Tibetan Plateau and pressing huge areas of Indochina and eastern China out to the east and southeast.

APPALACHIANS AND HIMALAYAS were formed by the same set of geologic processes, but roughly 250 million years apart. Many more years of erosion have given the older Appalachians (*left*) a less rugged appearance than the Himalayas (*right*), which are still being uplifted by strong tectonic forces.

Construction of the Himalayas and the Tibetan Plateau

ranges and up toward the surface. And in this manner, erosion leaves a distinct fingerprint on the rocks and on the pattern of crustal deformation in and under mountains.

The type of rock at the surface of a mountain is determined, in part, by the local climate and by the rate and pattern of erosion. In this way, erosion influences both the topography and the composition and structure of mountains. Metamorphism of rocks (changes as a result of heating and pressure) and the creation of many rock-forming minerals are governed by the pressure and temperature profile within the crust. Seemingly small details of climate and erosion, such as wind speed and direction or minor differences in latitude, can profoundly influence the temperature history, and therefore the type of rock created, as a mountain range evolves.

Computer models have examined the effects of prevailing wind direction and orography on the distribution of different metamorphic zones in mountain ranges. For mountains formed by subduction, prevailing winds in the same direction as subduction cause most of the precipitation to fall on the seaward side of the mountain range, which faces the subducting plate.

This phenomenon intensifies deformation and exhumation of rocks from deep in the crust. If, on the other hand, the prevailing winds are in the opposite direction as subduction, erosion is concentrated on the landward side of the mountain range, so that deformation is relatively uniform throughout the range and deep exhumation is limited to the interior, or continental, side of the range. One study of the eroded cores of several ancient mountain ranges revealed that the fingerprint of orography and wind direction remains clear, in the distribution of rocks sucked into the range by climatically driven erosion, up to two billion years after the ranges had become tectonically inactive.

With growing evidence that tectonic uplift and erosion can occur over similar timescales and at similar rates, many researchers have concluded that some mountain ranges have achieved a steady-state topography. In this state, the size of the mountains can remain stable for millions of years, because the rate of erosion matches the rate of uplift. Localized topography within such a mountain range will change as rocks of different strength are exposed at the surface. Average mountain height, however, may undergo little change, because of the long-term balance between tectonics and climate-driven erosion.

THREE STAGES

ALTHOUGH RELATIVELY FEW of Earth's mountains are now believed to be in perfect equilibrium, many of them may have achieved such a balance at some time in their history.

CLYDE H. SMITH *Peter Arnold, Inc.*

illustrates many of the principles of feedback-rich mountain building. For example, uplift of the plateau apparently triggered a climatic change around eight million years ago, which dramatically strengthened the Asian monsoon, the pattern of intense seasonal rainfall across southern Asia. The monsoon pattern sharply intensified erosion in the Himalayas, increasing the flux of sediment from the Indus and Bengal rivers by factors as high as 13. The strengthening of the Asian monsoon apparently caused a surge of uplift in the Himalayas, as the isostasy (buoyancy) of the crust responded to the intensified erosion in the region. Meanwhile the interior of the Tibetan Plateau evolved much more slowly because it lies in the rain shadow of the Himalayas and because the major rivers have not yet eroded their way into it.

Although the present-day Appalachians are less spectacular than the Himalayas, they were created by the same tectonic processes and are now being shaped by the same system feedback. The primary difference is age: the Himalayas are about 50 million years old, whereas the main uplift of the Appalachians culminated 250 million to 350 million years ago.

Geologically, the eastern coast of North America is the quiet side of the continent today. Before 200 million years ago, however, it was a hotbed of mountain building. During the previous several hundred million years, the predecessor to the Atlantic Ocean (called the Iapetus Ocean) was subducting underneath eastern North America. As the Iapetus gradually closed, at least three smaller landmasses, probably island arcs analogous to present-day Japan, slammed into the continent. Later, the mountain-building process culminated with the collision of Africa and the eastern U.S. The early Appalachians that resulted from these collisions are estimated to have been 250 to 350 kilometers wide, with average elevations of 3,500 to 4,500 meters and isolated peaks perhaps much higher. One study suggests that during the past 270 million years, erosion has stripped between 4,500 and 7,500 meters of material from the surface of the Appalachians. (This fact does not mean that the mountains were once 4,500 to 7,500 meters higher; isostatic uplift, as is explained in the main article, has continually pushed the roots of mountains upward in response to erosion at the surface.) Over the past 200 million years, as North America rifted away from Africa and the Atlantic Ocean began to open, secondary events may have triggered minor episodes of uplift, but erosion has been the dominant process shaping the mountain range. *—N.P. and M.T.B.*

Mountain ranges, it appears, often go through three distinct phases. The first, formative stage begins with the converging of plates or some other tectonic event that thickens crust and causes topography to rise. During this stage, rates of uplift exceed those of erosion. Erosion rates increase dramatically, however, as elevations and relief increase. Depending on the size of the range and the local climate, uplift may persist until erosion rates or the strength of the crust limits the average elevation of the range from increasing any more. This is the second stage, a steady state that may continue as long as the rates of uplift and erosion remain equal. When uplift diminishes, erosion begins to dominate and the final stage begins. In this final stage, the average elevation of the mountain range begins a long, slow decline. The cycle may be interrupted or complicated at any stage by tectonic or climatic events as well as the feedback among those processes and erosion.

The new model of how mountains develop promises to be as revolutionary as was plate tectonics some four decades ago. Just as plate tectonics managed to explain the worldwide distribution of earthquakes, volcanoes, fossils and many different rocks and minerals, the new understanding of mountain building shows how tectonic forces, Earth's climate and topography interact to create some of Earth's most spectacular landscapes. Like plate tectonics, the new model also illuminates phenomena that had long puzzled geologists. Computer simulations incorporating many of the model's principal precepts, for example, have proved very successful in mimicking the effects of complex tectonic histories, climatic variability and different geologic settings. Continuing research will provide even more details of how Earth's magnificent mountain ranges grow, evolve and decline, as well as details concerning the importance of mountains in shaping the climate and tectonics of our planet. SA

GALEN ROWELL *Mountain Light*

MORE TO EXPLORE

Landform Development by Tectonics and Denudation. T. Yoshikawa in *Themes in Geomorphology*. Edited by A. Pitty. Croom Helm, 1985.

Tectonic Forcing of Late Cenozoic Climate. M. E. Raymo and W. F. Ruddiman in *Nature,* Vol. 359, pages 117–122; September 10, 1992.

Climatic Limits on Landscape Development in the Northwestern Himalaya. N. Brozović, D. W. Burbank and A. J. Meigs in *Science,* Vol. 276, pages 571–574; April 25, 1997.

Landscape Evolution: The Interactions of Tectonics and Surface Processes. D. W. Burbank and N. Pinter in *Basin Research,* Vol. 11, pages 1–6; March 1999.

Orogeny and Orography: The Effects of Erosion on the Structure of Mountain Belts. S. D. Willett in *Journal of Geophysical Research,* Vol. 104, pages 28,957–28,981; December 1999.

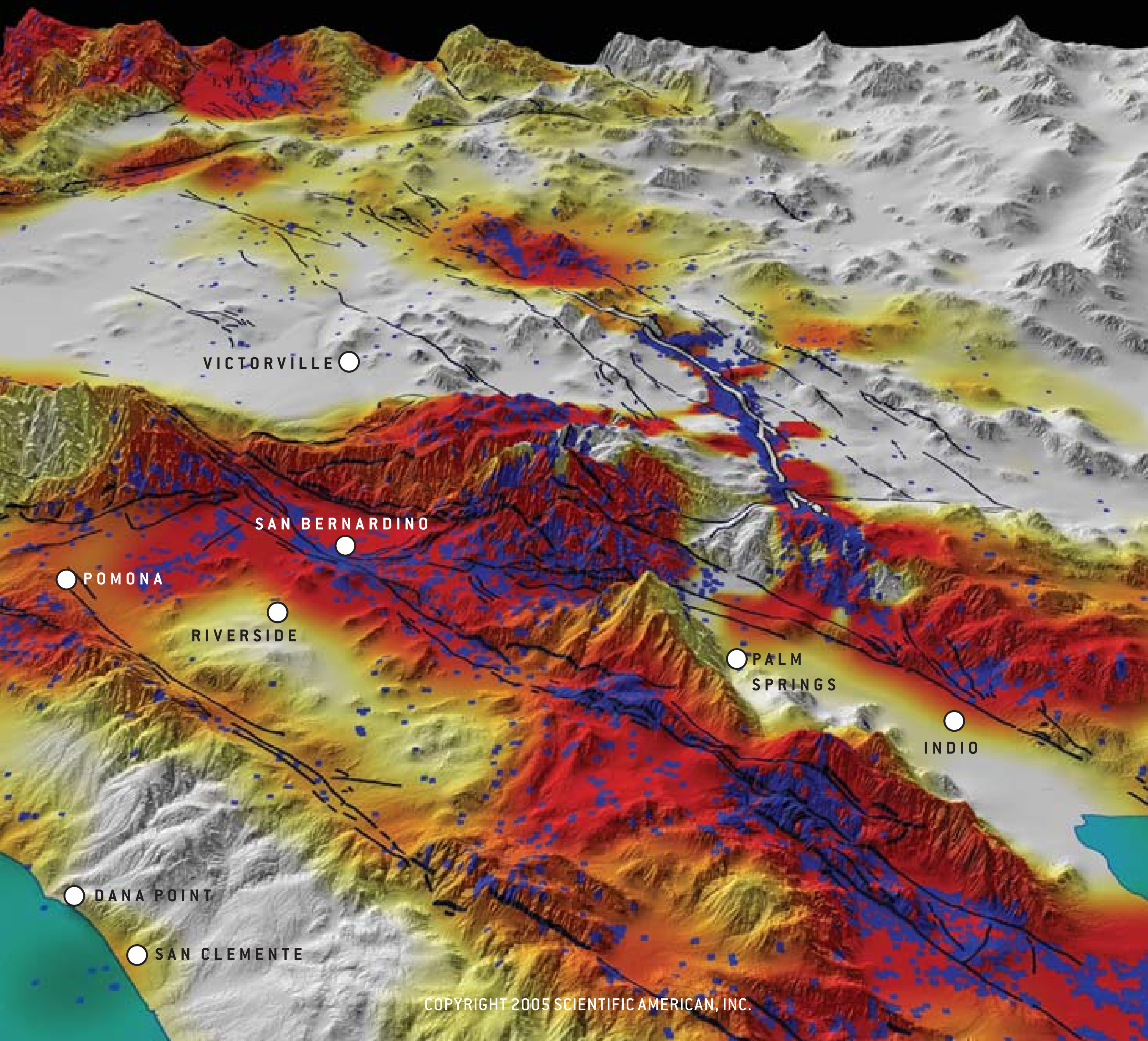
Mountain Building. E. A. Keller and N. Pinter in *Active Tectonics: Earthquakes, Uplift and Landscape.* Second edition. Prentice Hall, 2002.

Earthquake Conversations

Contrary to prevailing wisdom, large earthquakes can trigger or inhibit one another in unexpected ways. This exciting discovery could dramatically improve scientists' ability to pinpoint future shocks

By Ross S. Stein

THIRTEEN MILLION PEOPLE in and around Los Angeles co-exist with seismicity in the form of frequent earthquakes and tremors. Here the seismicity observed in southern California from 1996 to 1999 is represented by purple dots, with all shocks greater than or equal to magnitude 1.4 plotted. These actual tremors can be compared with the seismicity predicted by the time-dependent stress changes imparted by large earthquakes such as the 1992 magnitude 7.4 Landers earthquake (*visible as a white-inscribed black line in the center of the image*). Predicted seismicity is represented by the warm colors: the greater the predicted number of shocks, the redder the color.



For decades, earthquake experts dreamed of being able to divine the time and place of the world's next disastrous shock. But by the early 1990s the behavior of quake-prone faults had proved so complex that they were forced to conclude that the planet's largest tremors are isolated, random and utterly unpredictable. Most seismologists now assume that once a major earthquake and its expected aftershocks do their damage, the fault will remain quiet until stresses in Earth's crust have time to rebuild, typically over hundreds or thousands of years. A recent discovery—that earthquakes interact in ways never before imagined—is beginning to overturn that assumption.

This insight corroborates the idea that a major shock relieves stress—and thus the likelihood of a second major tremor—in some areas. But it also suggests that the probability of a succeeding earthquake elsewhere along the fault or on a nearby fault can actually jump by as much as a factor of three. To the people who must stand ready to provide emergency services or to those who set prices for insurance premiums, these refined predictions can be critical in determining which of their constituents are most vulnerable.

At the heart of this hypothesis—known as stress triggering—is the realization that faults are unexpectedly responsive to subtle stresses they acquire as neighboring faults shift and shake. Drawing on records of past tremors and novel calculations of fault behavior, my colleagues and I have learned that the stress relieved during an earthquake does not simply dissipate; instead it moves down the fault and concentrates in sites nearby. This jump in stress promotes subsequent tremors. Indeed, studies of about two dozen faults since 1992 have convinced many of us that earthquakes can be triggered even when the stress swells by as little as one eighth the pressure required to inflate a car tire.

Such subtle cause-and-effect relations among large shocks were not thought to exist—and never played into seismic forecasting—until recently. As a result, many scientists have been understandably skeptical about embracing this basis for a new approach to forecasting. Nevertheless, the stress-triggering hypothesis has continued to gain credibility through its ability to explain the location and frequency of earthquakes that followed several destructive shocks in California, Japan and Turkey. The hope of furnishing better warnings for such disasters is the primary motivation behind our ongoing quest to interpret these unexpected conversations between earthquakes.

THE AUTHOR

ROSS S. STEIN is a geophysicist with the U.S. Geological Survey's Earthquake Hazards Team in Menlo Park, Calif. He joined the survey in 1981 after earning a Ph.D. from Stanford University in 1980. Stein's research, which has been devoted to improving scientists' ability to assess earthquake hazards, has been funded by U.S. agencies such as the Office of Foreign Disaster Assistance and by private companies, including the European insurance company Swiss Re. For the work outlined in this article, Stein received the Eugene M. Shoemaker Distinguished Achievement Award of the USGS in 2000. He also presented the results in his Frontiers of Geophysics Lecture at the annual meeting of the American Geophysical Union in 2001. Stein has appeared in several TV documentaries, including *Great Quakes: Turkey* (Learning Channel, 2001) and the 2004 IMAX film *Forces of Nature*.

AFTERSHOCKS IGNORED

CONTRADICTING the nearly universal theory that major earthquakes strike at random was challenging from the start—especially considering that hundreds of scientists searched in vain for more than three decades to find predictable patterns in global earthquake activity, or seismicity. Some investigators looked for changing rates of small tremors or used sensitive instruments to measure Earth's crust as it tilts, stretches and migrates across distances invisible to the naked eye. Others tracked underground movements of gases, fluids and electromagnetic energy or monitored tiny cracks in the rocks to see whether they open or close before large shocks. No matter what the researchers examined, they found little consistency from one major earthquake to another.

Despite such disparities, historical records confirm that about one third of the world's recorded tremors—so-called aftershocks—cluster in space and time. All true aftershocks were thought to hit somewhere along the segment of the fault that slipped during the main shock. Their timing also follows a routine pattern, according to observations first made in 1894 by Japanese seismologist Fusakichi Omori and since developed into a basic principle known as Omori's law. Aftershocks are most abundant immediately after a main shock. Ten days later the rate of aftershocks drops to 10 percent of the initial rate, 100 days later it falls to 1 percent, and so on. This predictable jump and decay in seismicity means that an initial tremor modifies Earth's crust in ways that raise the prospect of succeeding ones, contradicting the view that earthquakes occur randomly in time. But because aftershocks are typically smaller than the most damaging quakes scientists would like to be able to predict, they were long overlooked as a key to unlocking the secrets of seismicity.

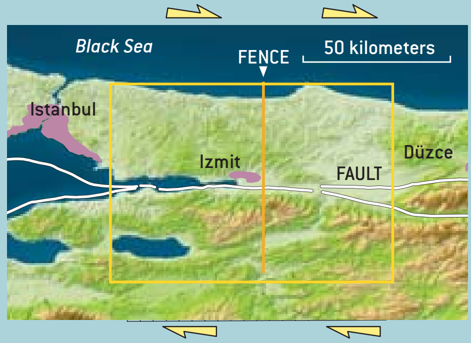
Once aftershocks are cast aside, the remaining tremors indeed appear—at least at first glance—to be random. But why ignore the most predictable earthquakes to prove that the rest are without order? My colleagues and I decided to hunt instead for what makes aftershocks so regular. We began our search in one of the world's most seismically active regions—the San Andreas Fault system that runs through California. From local records of earthquakes and aftershocks, we knew that on the day following a magnitude 7.3 event, the chance of another large shock striking within 100 kilometers is nearly 67 percent—20,000 times the likelihood on any other day. Something about the first shock dramatically increases the odds of subsequent ones, but what?

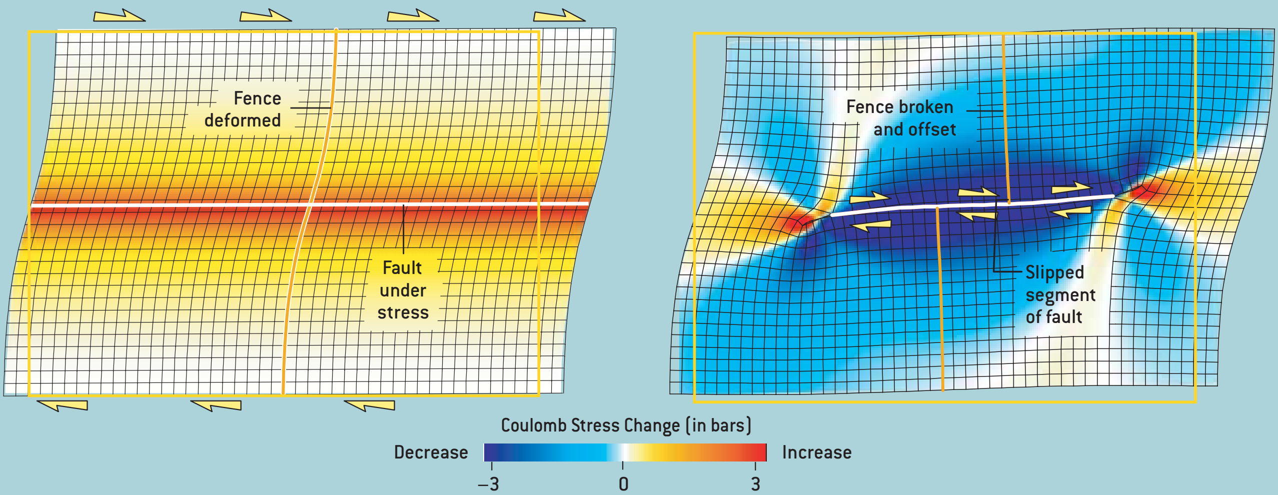
That big leap in probability explains why no one was ini-

SERKAN BOZKURT, KEITH RICHARDS-DINGER AND ROSS S. STEIN (*preceding pages*)

STRESSES AND SLIPS CAUSE DISASTER

BUILDUP AND RELEASE of the stress that accumulates slowly as Earth's tectonic plates grind past one another mark the cycle of all great earthquakes. Along Turkey's North Anatolian Fault (*white line*), the land north of the fault is moving eastward relative to the land to the south (*yellow arrows*) but gets stuck along the fault. When the stress finally overcomes friction along the fault, the rocks on either side slip past one another violently. A catastrophic example of this phenomenon occurred on August 17, 1999, when a magnitude 7.4 shock took 25,000 lives in and around the city of Izmit. Calculations of stress before and after the Izmit earthquake (*below*) reveal that, after the shock, the so-called Coulomb stress dropped along the segment of the fault that slipped but increased at the site of the subsequent magnitude 7.1 Düzce shock and also near Istanbul. —*R.S.S.*





BEFORE THE QUAKE

The segment of the North Anatolian Fault near Izmit accumulated significant stress (*red*) during the 200 years since its last major stress-relieving shock. An imaginary deformed fence and grid superimposed over the landscape illustrate this high stress (*warm colors*). Squares along the fault are stretched into parallelograms (*exaggerated 15,000 times*), with the greatest change in shape, and thus stress, occurring closest to the fault.

AFTER THE QUAKE

The earthquake relieved stress (*blue*) all along the segment of the fault that slipped. The formerly deformed fence broke and became offset by several meters at the fault, and the grid squares closest to the fault returned to their original shape. High stress (*warm colors*) is now concentrated beyond both ends of the failed fault segment, where the grid squares are more severely contorted than before the shock struck.

SERKAN BOZKURT *U.S. Geological Survey*

tially surprised in June 1992 when a magnitude 6.5 earthquake struck near the southern California town of Big Bear only three hours after a magnitude 7.3 shock occurred 40 kilometers away, near Landers. (Fortunately, both events took place in the sparsely populated desert and left Los Angeles unscathed.) But the puzzling contradiction to prevailing wisdom was that the Big Bear shock struck far from the fault that had slipped during Landers's shaking. Big Bear fit the profile of an aftershock in its timing but not in its location. We suspected that its mysterious placement might hold the clue we were looking for.

By mapping the locations of Landers, Big Bear and hundreds of other California earthquakes, my colleagues and I began to notice a remarkable pattern in the distribution not only of true aftershocks but also of other, smaller earthquakes that follow a main shock by days, weeks or even years. Like the enigmatic Big Bear event, a vast majority of these subsequent tremors tended to cluster in areas far from the fault that slipped during the earthquake and thus far from where aftershocks are supposed to occur [*see box on page 88*]. If we could determine what controlled this pattern, we reasoned, the same characteristics might also apply to the main quakes themselves. And if that turned out to be true, we might be well on our way to developing a new strategy for forecasting earthquakes.

TRIGGERS AND SHADOWS

WE BEGAN BY LOOKING at changes within Earth's crust after major earthquakes, which release some of the stress that accumulates slowly as the planet's shifting tectonic plates grind past one another. Along the San Andreas Fault, for instance, the plate carrying North America is moving south

FORECASTING UNDER STRESS

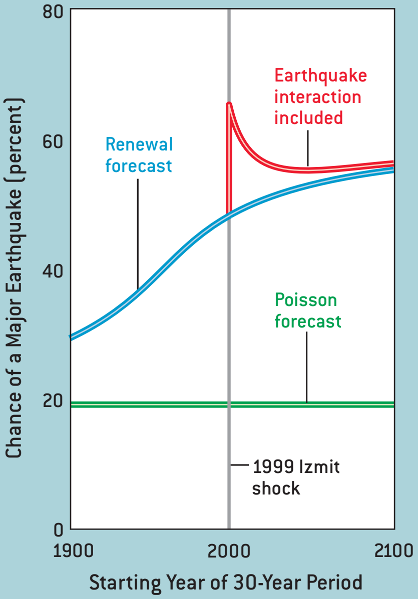
HOW PEOPLE PERCEIVE the threat of an earthquake in their part of the world depends in great part on what kind of warnings are presented to them. Most of today's seismic forecasts assume that one earthquake is unrelated to the next. Every fault segment is viewed as producing tremors of various sizes, each with an average frequency. For example, on average, magnitude 7 shocks might recur every 100 years and magnitude 6 shocks every 10 years. But the specific timing of the shocks is believed to be random. The best feature of this method, known as a Poisson probability, is that a forecast can be made without knowing when the last significant earthquake occurred. Seismologists can simply infer the typical time period between major shocks based on geologic records of much older tremors along that segment of the fault. This conservative strategy yields odds that do not change with time.

In contrast, a more refined type of forecast called the renewal probability predicts that the chances of a damaging shock climb as time passes since the last one struck. These growing odds are based on the assumption that stress along a fault increases gradually in the wake of a major earthquake. My colleagues and I build the probabilities associated with earthquake interactions on top of this second traditional technique by including the effects of stress changes imparted by nearby earthquakes. Comparing the three types of forecasts for Turkey's North Anatolian Fault near Istanbul illustrates their differences, which are most notable immediately after a major shock.

In the years leading up to the catastrophic Izmit earthquake of August 1999, the renewal probability of a shock of magnitude 7 or greater on the four faults within 50 kilometers of Istanbul had been rising slowly since the last large earthquake struck each of them, between 100 and 500 years ago. According to this type of forecast, the August 1999 shock created a sharp drop in the likelihood of a second major tremor in the immediate vicinity of Izmit, because the stress there is thought to have relaxed. But the quake caused no change in the 48 percent chance of severe shaking 100 kilometers to the west, in Istanbul, sometime in the next 30 years. Those odds will continue to grow slowly with time—unlike the Poisson probability, which will remain at only 20 percent regardless of other tremors that may occur near Istanbul.

When the effects of my team's new stress-triggering hypothesis were added to the renewal probability, everything changed. The most dramatic result was that the likelihood of a second quake rocking Istanbul shot up suddenly because some of the stress relieved near Izmit during the 1999 shock moved westward along the fault and concentrated closer to the city. That means the Izmit shock raised the probability of an Istanbul quake in the next 30 years from 48 percent to 62 percent. This so-called interaction probability will continue to decrease over time as the renewal probability climbs. The two forecasts will then converge at about 54 percent in the year 2060—assuming the next major earthquake does not occur before then. —*R.S.S.*

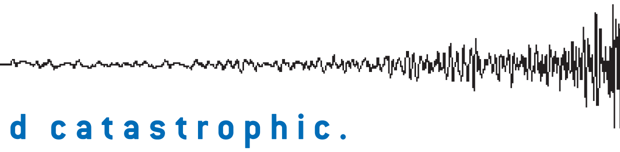
CRUMPLED BUILDINGS dot Düzce, Turkey, in the wake of a major tremor that struck in November 1999. Some scientists suspect that this disaster was triggered by an earlier shock near Izmit.



PREDICTED ODDS of a major earthquake striking within 50 kilometers of Istanbul can vary dramatically. The odds, which stay the same or rise slowly with time in traditional forecasts (*green* and *blue*), jump significantly when stresses transferred during the 1999 Izmit earthquake are included (*red*).

HURRIYET AP Photo (*photograph*); CORNELIA BLIK (*graph*)

Even tiny stress changes can have momentous effects, both calming and catastrophic.



relative to the one that underlies the Pacific Ocean. As the two sides move in opposite directions, shear stress is exerted parallel to the plane of the fault; as the rocks on opposite sides of the fault press against one another, they exert a second stress, perpendicular to the fault plane. When the shear stress exceeds the frictional resistance on the fault or when the stress pressing the two sides of the fault together is eased, the rocks on either side will slip past one another suddenly, releasing tremendous energy in the form of an earthquake. Both components of stress, which when added together are called Coulomb stress, diminish along the segment of the fault that slips. But because that stress cannot simply disappear, we knew it must be redistributed to other points along the same fault or to other faults nearby. We also suspected that this increase in Coulomb stress could be sufficient to trigger earthquakes at those new locations.

Geophysicists had been calculating Coulomb stresses for years, but scientists had never used them to explain seismicity. Their reasoning was simple: they assumed that the changes were too meager to make a difference. Indeed, the amount of stress transferred is generally quite small—less than 3.0 bars, or at most 10 percent of the total change in stress that faults typically experience during an earthquake. I had my doubts about whether this could ever be enough to trigger a fault to fail. But when Geoffrey King of the Paris Geophysical Institute, Jian Lin of the Woods Hole Oceanographic Institution in Massachusetts and I calculated the areas in southern California where stress had increased after major earthquakes, we were amazed to see that the increases—small though they were—matched clearly with sites where the succeeding tremors had clustered. The implications of this correlation were unmistakable: regions where the stress rises will harbor the majority of subsequent earthquakes, both large and small. We also began to see something equally astonishing: small reductions in stress could inhibit future tremors. On our maps, earthquake activity plummeted in these so-called stress shadows.

Coulomb stress analysis nicely explained the locations of certain earthquakes in the past, but a more important test would be to see whether we could use this new technique to forecast the sites of *future* earthquakes reliably. Eight years ago I joined geophysicist James H. Dieterich of the U.S. Geological Survey and geologist Aykut A. Barka, then at Istanbul Technical University, to assess Turkey's North Anatolian Fault, among the world's most heavily populated fault zones. Based on our calculations of where Coulomb stress had risen as a result of past earthquakes, we estimated that there was a 12 percent chance that a magnitude 7 shock or larger would strike the segment of the fault near the city of Izmit sometime between 1997 and 2027. That may seem like fairly low odds, but in comparison, all but one other segment of the 1,000-kilometer-long fault had calculated odds of only 1 to 2 percent.

We did not have to wait long for confirmation. In August 1999 a magnitude 7.4 quake devastated Izmit, killing 25,000 people and destroying more than $6.5 billion worth of property. But this earthquake was merely the most recent in a falling-domino-style sequence of 12 major shocks that had struck the North Anatolian Fault since 1939. In a particularly brutal five-year period, fully 700 kilometers of the fault slipped in a deadly westward progression of four shocks. We suspected that stress transferred beyond the end of each rupture triggered the successive earthquake, including Izmit's.

In November 1999 the 13th domino fell. Some of the Coulomb stress that had shifted away from the fault segment near Izmit triggered a magnitude 7.1 earthquake near the town of Düzce, about 100 kilometers to the east. Fortunately, Barka had calculated the stress increase resulting from the Izmit shock and had published it in the journal *Science* two months earlier. Barka's announcement had emboldened engineers to close school buildings in Düzce that were lightly damaged by the first shock despite pleas by school officials who said that students had nowhere else to gather for classes. Some of these buildings were flattened by the November shock.

If subsequent calculations by Parsons of the USGS, Shinji Toda of Japan's Active Fault Research Center, Barka, Dieterich and me are correct, that may not be the last of the Izmit quake's aftermath. The stress transferred during that shock has also raised the probability of strong shaking in nearby Istanbul over the next few years from 2 to 4 percent per year. Over the next 30 years we estimate the odds for a major quake to be 62 percent; if we assumed large shocks occur randomly, the odds would be just 20 percent [*see box on opposite page*].

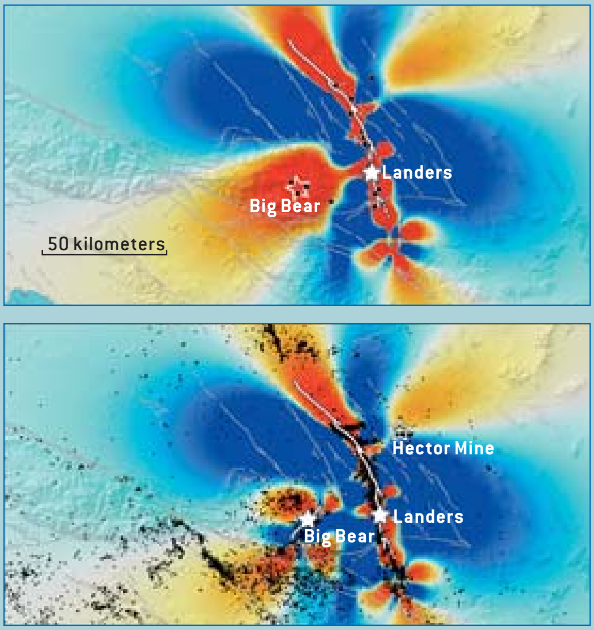
The stress-triggering hypothesis offers some comfort alongside such gloom and doom. When certain regions are put on high alert for earthquakes, the danger inevitably drops in others. In Turkey the regions of reduced concern happen to be sparsely populated relative to Istanbul. But occasionally the opposite is true. One of the most dramatic examples is the relative lack of seismicity that the San Francisco Bay Area, now home to six million people, has experienced since the great magnitude 7.9 earthquake of 1906. A 1998 analysis by my USGS colleagues Ruth A. Harris and Robert W. Simpson demonstrated that the stress shadows of the 1906 shock fell across several parallel strands of the San Andreas Fault in the Bay Area, whereas the stress increases occurred well to the north and south. This could explain why the rate of damaging shocks in the Bay Area dropped by an order of magnitude compared with the 75 years preceding 1906. Seismicity in the Bay Area is calculated to slowly emerge from this shadow as stress re-

UNIVERSITY OF SOUTHERN CALIFORNIA AND CONSORTIUM OF ORGANIZATIONS FOR STRONG-MOTION OBSERVATION SYSTEMS (*reversed seismogram of Whittier Narrows, Calif., earthquake*)

QUAKE CLUSTERS

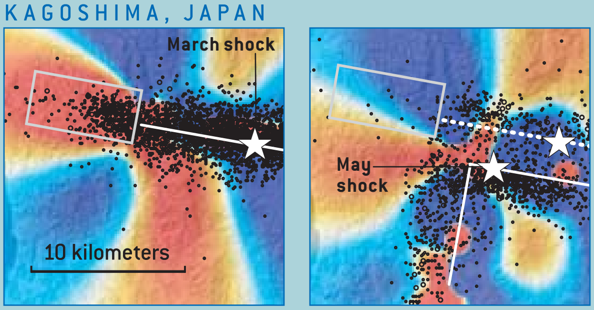
PLACES WHERE STRESS JUMPS (*red*) after major earthquakes (*filled stars*) tend to be the sites of subsequent tremors, both large (*open stars*) and small (*black dots*). Conversely, few tremors occur where the stress plummets (*blue*), regardless of the location of nearby faults (*white lines*). —*R.S.S.*

SOUTHERN CALIFORNIA, U.S.



MAGNITUDE 7.3 SHOCK in the southern California desert near Landers in 1992 increased the expected rate of earthquakes to the southwest, where the magnitude 6.5 Big Bear shock struck three hours later (*top*). Stresses imparted by the combination of the Landers and Big Bear events coincided with the regions where the vast majority of tremors occurred over the next seven years, culminating with the magnitude 7.1 Hector Mine quake in 1999 (*bottom*).

KAGOSHIMA, JAPAN



TWIN EARTHQUAKES can turn the rate of earthquake occurrence, or seismicity, up and down in the same spot. In March 1997 a magnitude 6.5 tremor increased stress and seismicity to the west of the ruptured fault (*above left*). Seismicity in that area then dropped along with stress (*above right*) following a magnitude 6.3 shock that struck 48 days later three kilometers to the south.

builds on the faults; the collapsed highways and other damage wrought by the 1989 Loma Prieta shock—as well as a flurry of small shocks since 2000 along the Hayward, Rodgers Creek and Calaveras faults—may be a harbinger of this reawakening.

BOLSTERING THE HYPOTHESIS

EXAMINATIONS OF THE EARTHQUAKES in Turkey and in southern California fortified our assertion that even tiny stress changes can have momentous effects, both calming and catastrophic. But despite the growing number of examples we had to support this idea, one key point was difficult to explain: roughly one quarter of the earthquakes we examined occurred in areas where stress had *decreased*. It was easy for our more skeptical colleagues to argue that no seismicity should occur in these shadow zones, because the main shock would have relieved at least some stress and thus pushed those segments of the fault further from failure. We now have an answer. Seismicity never shuts off completely in the shadow zones, nor does it turn on completely in the trigger zones. Instead the *rate* of seismicity—the number of earthquakes per unit of time—merely drops in the shadows or climbs in the trigger zones relative to the preceding rate in that area.

We owe this persuasive extension of stress triggering to a theory proposed by Dieterich in 1994. Known as rate/state friction, it jettisons the comfortable concept of friction as a property that can vary only between two values—high friction when the material is stationary and lower friction when it is sliding. Rather faults can become stickier or more slippery as the rate of movement along the fault changes and as the history of motion, or the state, evolves. These conclusions grew out of lab experiments in which Dieterich's team sawed a miniature fault into a Volkswagen-size slab of granite and triggered tiny earthquakes.

When earthquake behavior is calculated with friction as a variable rather than a fixed value, it becomes clear that Omori's law is a fundamental property not just of so-called aftershocks but of *all* earthquakes. The law's prediction that the rate of shocks will first jump and then diminish with time explains why a region does not forever retain the higher rate of seismicity that results from an increase in stress. But that is only half the story. Dieterich's theory reveals a characteristic of the seismicity that Omori's law misses entirely. In areas where a main shock relieves stress, the rate of seismicity immediately plunges but will slowly return to preshock values in a predictable manner. These points may seem subtle, but rate/state friction allowed us for the first time to make predictions of how jumps or declines in seismicity would change over time. When calculating Coulomb stresses alone, we could define the general location of new earthquakes but not their timing.

Our emerging ideas about both the place and the time of stress-triggered earthquakes were further confirmed by a global study published in 2005. Parsons considered the more than 100 earthquakes of magnitude 7 or greater that have occurred worldwide in the past 25 years and then ex-

SERKAN BOZKURT *USGS*, KEITH RICHARDS-DINGER *U.S. Navy* AND SHINJI TODA *Japan's Active Fault Research Center*

Taking stress triggering into account will raise different faults to the top of the high-alert list.

CALIFORNIA STRONG MOTION INSTRUMENTATION PROGRAM AND CONSORTIUM OF ORGANIZATIONS FOR STRONG-MOTION OBSERVATION SYSTEMS (*seismogram of Hector Mine, Calif., earthquake*)

amined all subsequent shocks of at least magnitude 5 within 250 kilometers of each magnitude 7 event. Among the more than 2,000 shocks in this inventory, 61 percent occurred at sites where a preceding shock increased the stress, even by a small amount. Few of these triggered shocks were close enough to the main earthquake to be considered an aftershock, and in all instances the rate of these triggered tremors decreased in the time period predicted by rate/state friction and Omori's law.

Now that we are regularly incorporating the concept of rate/state friction into our earthquake analyses, we have begun to uncover more sophisticated examples of earthquake interaction than Coulomb stress analyses alone could have illuminated. Until recently, we had explained only relatively simple situations, such as those in California and Turkey, in which a large earthquake spurs seismicity in some areas and makes it sluggish in others. We knew that a more compelling case for the stress-triggering hypothesis would be an example in which successive, similar-size shocks are seen to turn the frequency of earthquakes up and down in the same spot, like a dimmer switch on an electric light.

Toda and I discovered a spectacular example of this phenomenon, which we call toggling seismicity. In 2002 we began analyzing an unusual pair of magnitude 6.5 earthquakes that struck Kagoshima, Japan, in 1997. Immediately following the first earthquake, which occurred in March, a sudden burst of seismicity cropped up in a 25-square-kilometer region just beyond the west end of the failed segment of the fault. When we calculated where the initial earthquake transferred stress, we found that it fell within the same zone as the heightened seismicity. We also found that the rate immediately began decaying just as rate/state friction predicted. But when the second shock struck three kilometers to the south only seven weeks later, the region of heightened seismicity experienced a sudden, additional drop of more than 85 percent. In this case, the trigger zone of the first earthquake had fallen into the shadow zone of the second one. In other words, the first quake turned seismicity up, and the second one turned it back down.

A NEW GENERATION OF FORECASTS

EAVESDROPPING ON the conversations between earthquakes has revealed, if nothing else, that seismicity is highly interactive. And although phenomena other than stress transfer may influence these interactions, my colleagues and I believe that enough evidence exists to warrant an overhaul of traditional probabilistic earthquake forecasts. By refining the likelihood of dangerous tremors to reflect subtle jumps and declines in stress, these new assessments will help governments, the insurance industry and the public at large to better evaluate their earthquake risk. Traditional strategies already make some degree of prioritizing possible, driving the strengthening of buildings and other precautions in certain cities or regions at the expense of others. But our analyses have shown that taking stress triggering into account will raise different faults to the top of the high-alert list than using traditional methods alone will. By the same token, a fault deemed dangerous by traditional practice may actually be a much lower risk.

Regardless of the factors considered, chance plays a tremendous role in whether a large earthquake occurs, just as it does in whether a particular weather pattern produces a rainstorm. The meteorologists' advantage over earthquake scientists is that they have acquired millions more of the key measurements that help improve their predictions.

That is why my team is building an inventory of forecasts for large earthquakes near the shock-prone cities of Istanbul, Landers, San Francisco and Kobe. We are also making assessments for southern California [*see illustration on page 83*] and Tokyo, where a major earthquake could wreak a great loss of lives and trillion-dollar devastation.

The Indonesian earthquake of December 26, 2004, reminds us that a great earthquake's death toll can extend well beyond the reach of strong shaking when it triggers a large tsunami. This disastrous earthquake also left a seismic legacy: large Coulomb stress increases along Indonesia's Sumatra Fault and the parallel Sunda trench. Illustrating how the stress-triggering hypothesis can help in earthquake forecasting, my colleagues John McCloskey, Suleyman S. Nalbant and Sandy Steacy of the University of Ulster in Northern Ireland warned in the March 17, 2005, issue of *Nature* that these stresses could trigger another great earthquake on either of these faults. Just 11 days later a magnitude 8.7 earthquake occurred on the Sunda trench.

In the end, the degree to which any probabilistic forecast will protect people and property is still uncertain. But scientists have plenty of reasons to keep pursuing this dream: several hundred million people live and work along the world's most active fault zones. With that much at stake, stress triggering—or any other phenomenon that has the potential to raise the odds of a damaging earthquake—should not be ignored. SA

MORE TO EXPLORE

The Role of Stress Transfer in Earthquake Occurrence. R. S. Stein in *Nature*, Vol. 402, pages 605–609; December 9, 1999.

Heightened Odds of Large Earthquakes Near Istanbul: An Interaction-Based Probability Calculation. Tom Parsons, Shinji Toda, Ross S. Stein, Aykut Barka and James H. Dieterich in *Science*, Vol. 288, pages 661–665, April 28, 2000. **http://quake.wr.usgs.gov/research/deformation/modeling/papers/istanbul.html**

View earthquake animations and download Coulomb 2.2 (software and tutorial for calculating earthquake stress changes) at **http://quake.usgs.gov/~ross**

The Threat of *Silent* Earthquakes

A lack of rumbling does not necessarily make an earthquake harmless. Some of the quiet types could presage devastating tsunamis or larger, ground-shaking shocks

By Peter Cervelli

In early November 2000 the Big Island of Hawaii experienced its largest earthquake in more than a decade. Some 2,000 cubic kilometers of the southern slope of Kilauea volcano lurched toward the ocean, releasing the energy of a magnitude 5.7 shock. Part of that motion took place under an area where thousands of people stop every day to catch a glimpse of one of the island's most spectacular lava flows. Yet when the earthquake struck, no one noticed—not even seismologists.

How could such a notable event be overlooked? As it turns out, quaking is not an intrinsic part of all earthquakes. The event on Kilauea was one of the first unambiguous records of a so-called silent earthquake, a type of massive earth movement unknown to science until just a few years ago. Indeed, I would never have discovered this quake if my colleagues at the U.S. Geological Survey's Hawaiian Volcano Observatory had not already been using a network of sensitive instruments to monitor the volcano's activity. When I finally noticed that Kilauea's south flank had shifted 10 centimeters along an underground fault, I also saw that this movement had taken nearly 36 hours—a turtle's pace for an earthquake. In a typical tremor, opposite sides of the fault rocket past each other in a matter of seconds—quickly enough to create the seismic waves that cause the ground to rumble and shake.

But just because an earthquake happens slowly and quietly does not make it insignificant. My co-investigators and I realized immediately that Kilauea's silent earthquake could be a harbinger of disaster. If that same large body of rock and debris were to gain momentum and take the form of a gigantic landslide—separating itself from the rest of the volcano and sliding rapidly into the sea—the consequences would be devastating. The collapsing material would push seawater into towering tsunami waves that could threaten coastal cities along

GIANT LANDSLIDE (*upper left*) spawned by a silent earthquake could generate a fearsome tsunami hundreds of meters high (*below*).

the entire Pacific Rim. Such catastrophic flank failure, as geologists call it, is a potential threat around many island volcanoes worldwide.

UNEXPECTED STIR

FORTUNATELY, the discovery of silent earthquakes is revealing more good news than bad. The chances of catastrophic flank failure are slim, and the instruments that record silent earthquakes might make early warnings possible. New evidence for conditions that might trigger silent slip suggests bold strategies for preventing flank collapse. Occurrences of silent earthquakes are also being reported in areas where flank failure is not an issue. There silent earthquakes are inspiring ways to improve forecasts of their ground-shaking counterparts.

The discovery of silent earthquakes and their link to catastrophic flank collapse was a by-product of efforts to study other potential natural hazards. Destructive earthquakes and volcanoes are a concern in Japan and the U.S. Pacific Northwest, where tectonic plates constantly plunge deep into the planet along what are called subduction zones. Beginning in the early 1990s, geologists began deploying large networks of continuously recording Global Positioning System (GPS) receivers in these regions and along the slopes of active volcanoes such as Kilauea. By receiving signals from a constellation of more than 30 navigational satellites, these instruments can measure their own positions on Earth's surface at any given time to within a few millimeters.

The scientists who deployed these GPS receivers expected to see both the slow, relentless motion of the planet's shell of tectonic plates and the relatively quick movements that earthquakes and volcanoes trigger. It came as some surprise when these instruments detected small ground movements that were not associated with any known earthquake or eruption. When researchers plotted the ground movements on a map, the pattern that resulted very much resembled one characteristic of fault movement. In other words, all the GPS stations on one side of a given fault moved several centimeters in the same general direction. This pattern would have been no surprise if it had taken a year or longer to form. In that case, scientists would have known that a slow and steady process called fault creep was responsible. But at rates of up to centimeters a day, the mystery events were hundreds of times faster than that. Beyond their relative speediness, these silent earthquakes shared another attribute with their noisy counterparts that distinguished them from fault creep: they are not steady processes but instead are discrete events that begin and end suddenly.

That sudden beginning, when it takes place on the slopes of a volcanic island, creates concern about a possible catastrophic flank event. Most typical earthquakes happen along faults that have built-in brakes: motion stops once the stress is relieved between the two chunks of earth that are trying to move past each other. But activity may not stop if gravity becomes the primary driver. In the worst-case scenario, the section of the volcano lying above the fault becomes so unstable that once slip starts, gravity pulls the entire mountainside downhill until it disintegrates into a pile of debris on the ocean floor.

The slopes of volcanoes such as Kilauea become steep and vulnerable to this kind of collapse when the lava from repeated eruptions builds them up more rapidly than they can erode away. Discovering the silent earthquake on Kilauea suggests that the volcano's south flank is on the move—perhaps on its way to eventual obliteration.

For now, friction along the fault is acting like an emergency brake. But gravity has won out in many other instances in the past. Scientists have long seen evidence of ancient collapses in sonar images of giant debris fields in the shallow waters surrounding volcanic islands around the world, including Majorca in the Mediterranean Sea and the Canary Islands in the Atlantic Ocean. In the Hawaiian Islands, geologists have found more than 25 individual collapses that have occurred over the past five million years—the blink of an eye in geologic time.

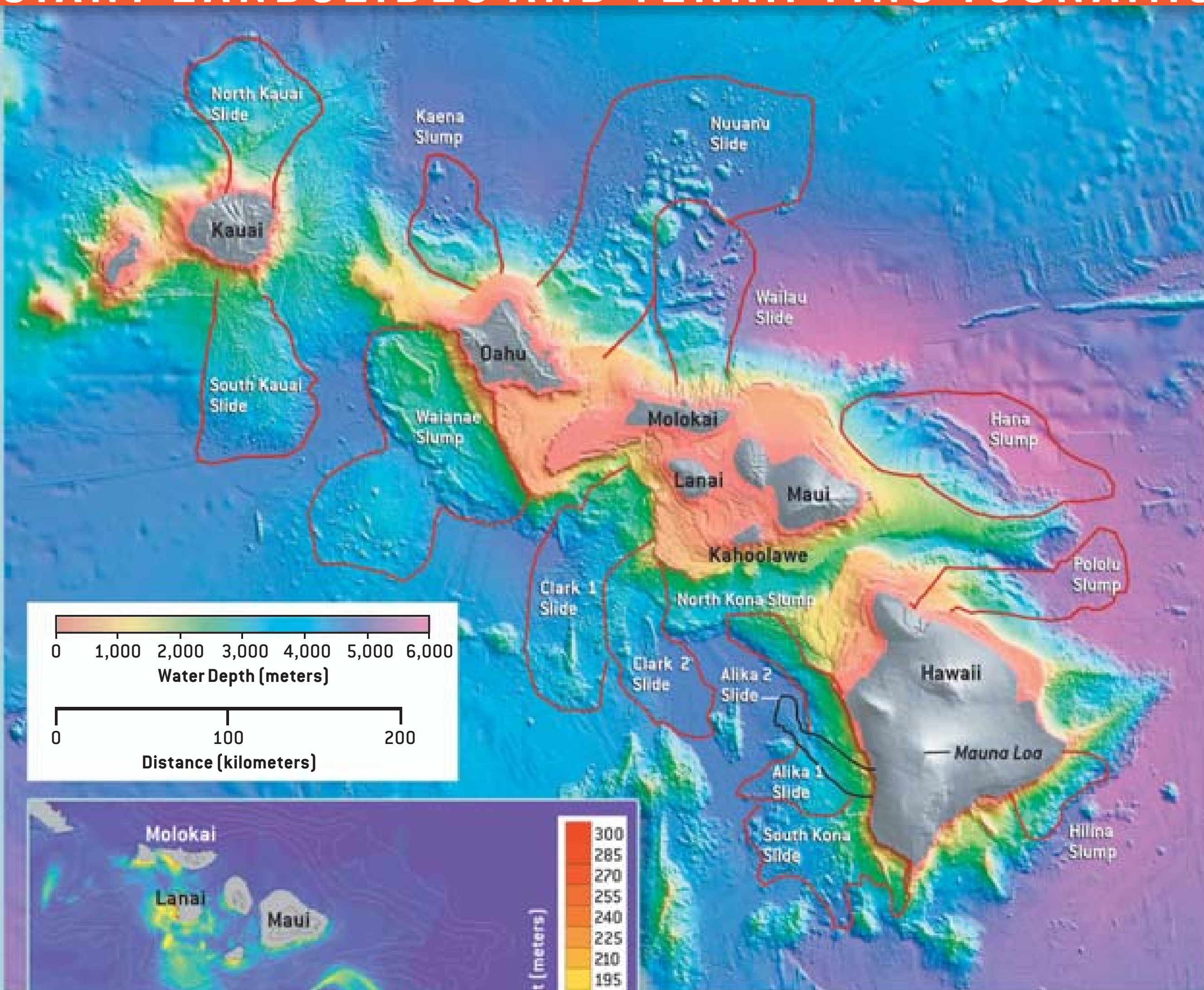
In a typical slide, the volume of material that enters the ocean is hundreds of times as great as the section of Mount St. Helens that blew apart during the 1980 eruption—more than enough to have triggered immense tsunamis. On the Hawaiian island of Lanai, for instance, geologists discovered evidence of wave action, including abundant marine shell fragments, at elevations of 325 meters. Gary M. McMurtry of the University of Hawaii at Manoa and his colleagues conclude that the most likely way the shells could have reached such a lofty location was within the waves of a tsunami that attained the astonishing height of 300 meters along some Hawaiian coastlines. Most of the tallest waves recorded in modern times were no more than one tenth that size. Indeed, the December 26, 2004, Indian

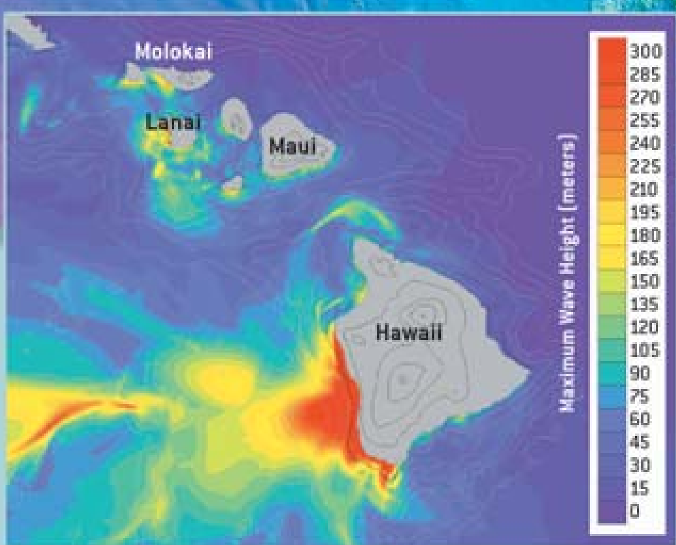
Overview/*Slippery Slope*

- Not all earthquakes shake the ground, it turns out. The so-called silent types are forcing scientists to rethink their understanding of the way quake-prone faults behave.
- In rare instances, silent earthquakes that occur along the flanks of seaside volcanoes may cascade into monstrous landslides that crash into the sea and trigger towering tsunamis.
- Silent earthquakes that take place within fault zones created by one tectonic plate diving under another may increase the chance of ground-shaking shocks.
- In other locations, however, silent slip may decrease the likelihood of destructive quakes, because they release stress along faults that might otherwise seem ready to snap.

DAVID FIERSTEIN (*preceding pages*)

GIANT LANDSLIDES AND TERRIFYING TSUNAMIS





UNDERWATER FIELDS of rock and debris (*examples outlined in red*) reveal that massive sections of Hawaiian volcanoes have crumbled into the ocean many times in the past. Some geologists suspect that one collapse (*black outline*) off the western flank of Mauna Loa volcano kicked up a gigantic tsunami that deposited shattered shell and rock as high as 800 meters along the nearby coast. A computer simulation (*left*) reveals that the same landslide may have produced waves up to 300 meters high. For a brief time, Maui could have been divided in two and the ocean floor west of Molokai could have been exposed.

USGS, WITH JAPAN MARINE SCIENCE AND TECHNOLOGY CENTER, UNIVERSITY OF HAWAII AND MONTEREY BAY AQUARIUM RESEARCH INSTITUTE (*top*); REPRINTED FROM GARY M. MCMURTRY IN *MARINE GEOLOGY*, VOL. 202, © 2002, WITH PERMISSION FROM ELSEVIER (*bottom*)

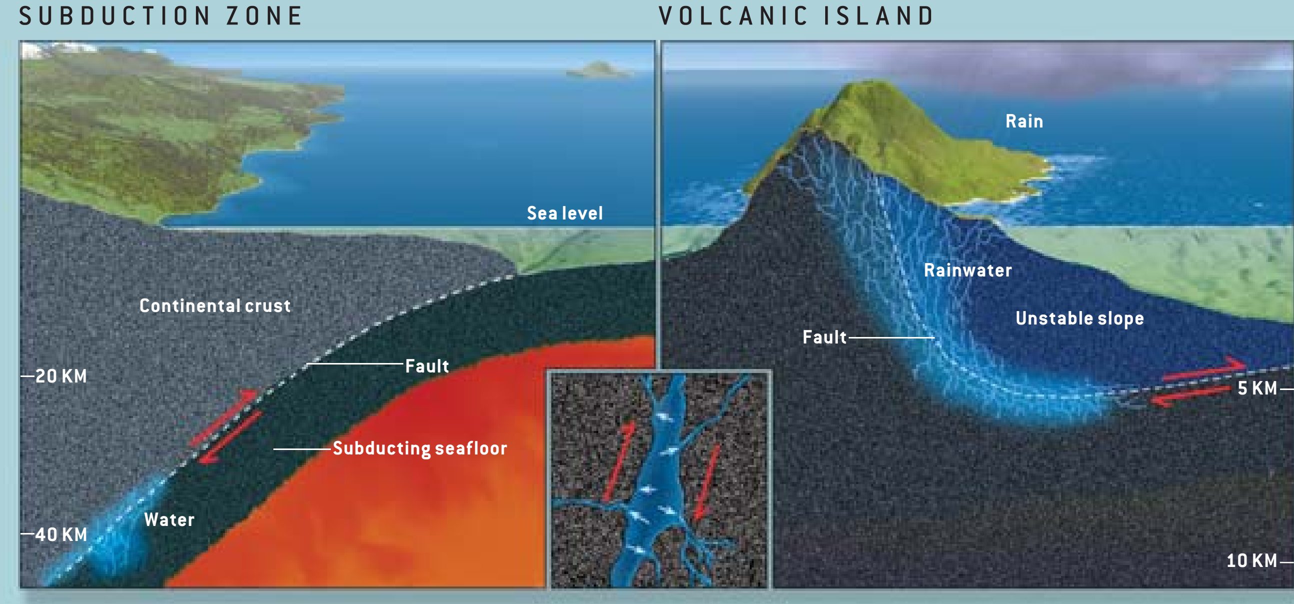
Ocean tsunami, which killed more than 200,000 people, reached "only" about 20 meters.

As frightening as such an event may sound, this hazard must be understood in the proper context. Catastrophic failure of volcanic slopes is very rare on a human timescale—though far more common than the potential for a large asteroid or comet to have a damaging collision with Earth. Collapses large enough to generate a tsunami occur somewhere in the Hawaiian Islands only about once every 100,000 years. Some scientists estimate that such events occur worldwide once every 10,000 years. Because the hazard is extremely destructive when it does happen, many scientists agree that it is worth preparing for.

To detect deformation within unstable volcanic islands, networks of continuous GPS receivers are beginning to be deployed on Réunion Island in the

THE MECHANICS OF SILENT EARTHQUAKES

PERCOLATING WATER may trigger silent earthquakes if it finds a way into a vulnerable fault. Highly pressurized by the burden of overlying rock, water can push apart the two sides of the fault (*inset*), making it easier for them to slip past each other (*red arrows*). This kind of silent slip can occur within subduction zones and volcanic islands.



Water-filled fault

WATER squeezed out of hydrous minerals in a slab of ancient seafloor may enter faults created as the slab dives underneath another tectonic plate.

RAINWATER may seep down from Earth's surface into shallow faults, such as those that separate an unstable slope from the rest of a volcano.

Indian Ocean, on Fogo in the Cape Verde Islands, and throughout the Galápagos archipelago, among others. Kilauea's network of more than 20 GPS stations, for example, has already revealed that the volcano experiences fault creep, silent earthquakes as well as large, destructive typical earthquakes. Some scientists propose, however, that Kilauea may currently be protected from catastrophic collapse by several underwater piles of mud and rock—probably debris from old flank collapses—that are buttressing its south flank. New discoveries about the way Kilauea is slipping can be easily generalized to other island volcanoes that may not have similar buttressing structures.

Whatever the specific circumstances for an island, the transition from silent slip to abrupt collapse would involve a sudden acceleration of the mobile slope. In the worst case, this acceleration would proceed immediately to breakneck velocities, leaving no chance for early detection and warning. In the best case, the acceleration would occur in fits and starts, in a cascade of silent earthquakes slowly escalating into regular earthquakes, and then on to catastrophe. A continuous GPS network could easily detect this fitful acceleration, well before ground-shaking earthquakes began to occur and, with luck, in plenty of time for a useful tsunami warning.

If the collapse were big enough, however, a few hours' or even days' warning might come as little comfort because it would be so difficult at that point to evacuate everyone. This problem raises the question of whether authorities might ever implement preventive measures. The problem of stabilizing the teetering flanks of oceanic volcanoes is solvable—in principle. In practice, however, the effort required would be immense. Consider simple brute force. If enough rock were removed from the upper reaches of an unstable volcanic flank, then the gravitational potential energy that is driving the system toward collapse would disappear for at least several hundred thousand years. A second possible method—lowering an unstable flank slowly through a series of small earthquakes—would be much cheaper but fraught with geologic unknowns and potential dangers. To do so, scientists could conceivably harness as a tool to prevent collapse the very thing that may be currently driving silent earthquakes on Kilauea.

THE AUTHOR

PETER CERVELLI is a research geophysicist at the U.S. Geological Survey's Alaska Volcano Observatory and affiliate professor at the Geophysical Institute of the University of Alaska–Fairbanks. He previously headed the Hawaiian Volcano Observatory's crustal deformation project. Cervelli discovered the silent earthquake that struck Kilauea's south flank in November 2000 while he was working on his Ph.D., which he received from Stanford University in 2001.

DAVID FIERSTEIN

Nine days before the most recent silent earthquake on Kilauea, a torrential rainstorm dropped nearly a meter of water on the volcano in less than 36 hours. Geologists have long known that water leaking into faults can trigger earthquakes, and nine days is about the same amount of time that they estimate it takes water to work its way down through cracks and pores in Kilauea's fractured basaltic rock to a depth of five kilometers—where the silent earthquake occurred. My colleagues and I suspect that the burden of the overlying rock pressurized the rainwater, forcing the sides of the fault apart and making it much easier for them to slip past each other.

This discovery lends credence to the controversial idea of forcefully injecting water or steam into faults at the base of an unstable flank to trigger the stress-relieving earthquakes needed to let it down slowly. This kind of human-induced slip happens at very small scales all the time at geothermal plants and other locations where water is pumped into Earth. But when it comes to volcanoes, the extreme difficulty lies in putting the right amount of fluid in the right place so as not to inadvertently generate the very collapse that is meant to be avoided. Some geophysicists considered this strategy as a way to relieve stress along California's infamous San Andreas fault, but they ultimately abandoned the idea for fear that it would create more problems than it would solve.

WEDGES OF WATER

APART FROM CALLING attention to the phenomenon of catastrophic collapse of the flank of a volcano, the discovery of silent earthquakes is prompting scientists to reconsider various aspects of fault motion—including seismic hazard assessments. In the U.S. Pacific Northwest, investigators have observed many silent earthquakes along the enormous Cascadia fault zone between the North American plate and the subducting Juan de Fuca plate. One curious feature of these silent earthquakes is that they happen at regular intervals—so regular, in fact, that scientists are now predicting their occurrence successfully.

This predictability most likely stems from the fact that water flowing from below subduction zones may exert significant control over when and where these faults slip silently. As the subducting plate sinks deeper into Earth, it encounters higher and higher temperatures and pressures, which release the significant amount of water trapped in water-rich minerals that exist within the slab. The silent earthquakes may then take place when a batch of fluid from the slab is working its way up—as the fluid passes, it will unclamp the fault zone a little bit, perhaps allowing some slow slip.

What is more, Garry Rogers and Herb Dragert of the Geological Survey of Canada reported in June 2003 that these silent tremors might even serve as precursors to some of the region's large, ground-shaking shocks. Because the slow slips occur deep and at discrete intervals, they regulate the rate at which stress accumulates on the shallower part of the fault zone, which moves in fits and starts. In this shallow, locked segment of the fault, it usually takes years or even centuries to amass the stress required to set off a major shock. Rogers and Dragert suggest, however, that silent slip may dramatically hasten this stress buildup, thereby increasing the risk of a regular earthquake in the weeks and months after a silent one.

Silent earthquakes are forcing scientists to rethink seismic forecasts in other parts of the world as well. Regions of Japan near several so-called seismic gaps—areas where fewer than expected regular earthquakes occur in an otherwise seismically active region—are thought to be overdue for a destructive shock. But if silent slip has been relieving stress along these faults without scientists realizing it, then the degree of danger may actually be less than they think. Likewise, if silent slip is discovered along faults that were considered inactive until now, these structures will need careful evaluation to determine whether they are also capable of destructive earthquakes.

RETHINKING DOCTRINES

IF FUTURE STUDY reveals silent earthquakes to be a common feature of most large faults, then scientists must revisit long-held doctrines about all earthquakes. The observation of many different speeds of fault slip poses a real challenge to theorists trying to explain the faulting process with fundamental physical laws, for example. It is now believed that the number and sizes of observed earthquakes can be explained with a fairly simple friction law. But can this law also account for silent earthquakes? So far no definitive answer has been found, but research continues.

Silent earthquakes are only just beginning to enter the public lexicon. These subtle events portend an exponential increase in our understanding of the how and why of fault slip. The importance of deciphering fault slip is difficult to overstate because when faults slip quickly, they can cause immense damage, sometimes at a great distance from the source. The existence of silent earthquakes gives scientists a completely new angle on the slip process by permitting the detailed study of fault zones through every stage of their movement. SA

MORE TO EXPLORE

Sudden Aseismic Fault Slip on the South Flank of Kilauea Volcano, Hawaii. Peter Cervelli, Paul Segall, Kaj Johnson, Michael Lisowski and Asta Miklius in *Nature,* Vol. 415, pages 1014–1017; February 28, 2002.

Episodic Tremor and Slip on the Cascadia Subduction Zone: The Chatter of Silent Slip. Garry Rogers and Herb Dragert in *Science,* Vol. 300, pages 1942–1943; June 20, 2003.

Giant Landslides, Mega-Tsunamis, and Paleo-Sea Level in the Hawaiian Islands. G. M. McMurtry, P. Watts, G. J. Fryer, J. R. Smith and F. Imamura in *Marine Geology.* Available at **www.sciencedirect.com/science/journal/00253227**

Visit the U.S. Geological Survey Hawaiian Volcano Observatory at **http://hvo.wr.usgs.gov**

New Light on Deep Earthquakes

Until about 15 years ago, it was a mystery how deep earthquakes could occur. Recent results have now demonstrated mechanisms for such rock failures at great depths

By Harry W. Green II

ON JUNE 7, 1994, a great earthquake rumbled through Earth's mantle more than 600 kilometers beneath Bolivia. It was the largest earthquake ever recorded at such depths. The tremors were felt as far away as Toronto. No other quake in history had been felt so far from its epicenter.

The event was truly spectacular and yet paradoxical as well. Although deep earthquakes are as regular as clockwork, in theory they should be impossible. Indeed, the very existence of deep earthquakes has teased geophysicists since their discovery in 1927. In 1989 my colleagues and I began to unravel the solution to this puzzle in the laboratory. This article gives an account of that discovery, the new theory of earthquakes that has flowed from it, and recent developments in our understanding.

Most earthquakes are shallow events that occur within a few tens of kilometers of Earth's surface by the familiar processes of brittle fracture and frictional sliding—the same mechanisms by which glass breaks and tires squeal on pavement. For example, last year's devastating Indonesian earthquake was a shallow quake only 16 to 18 kilometers beneath the Indian Ocean. But almost 30 percent of all earthquakes originate much deeper, occurring at depths exceeding 70 kilometers, where the pressure reaches upwards of two gigapascals (20,000 times that of the atmosphere at sea level); nearly 8 percent happen at depths greater than 300 kilometers, where the pressure is greater than 10 gigapascals. At such high pressures, rock will flow at lower stresses than those at which it will break or slide along a preexisting fault. Earthquakes at depth then, would seem impossible.

Nevertheless, deep earthquakes do occur, exclusively in thin, planar zones that begin beneath oceanic trenches and angle down into the mantle. The theory of plate tectonics tells us that these locations mark subduction zones, where the cold uppermost layer (the lithosphere, 50 to 100 kilometers thick) sinks into the mantle. In doing so, it provides the return flow that compensates for the upwelling of new lithosphere at ocean ridges. In these zones, earthquakes show an exponential decrease in frequency from the surface to about 300 kilometers deep. Then their frequency increases again, peaking at 550 to 600 kilometers deep. Finally, earthquakes cease entirely at approximately 680 kilometers deep.

Elusive Mechanism

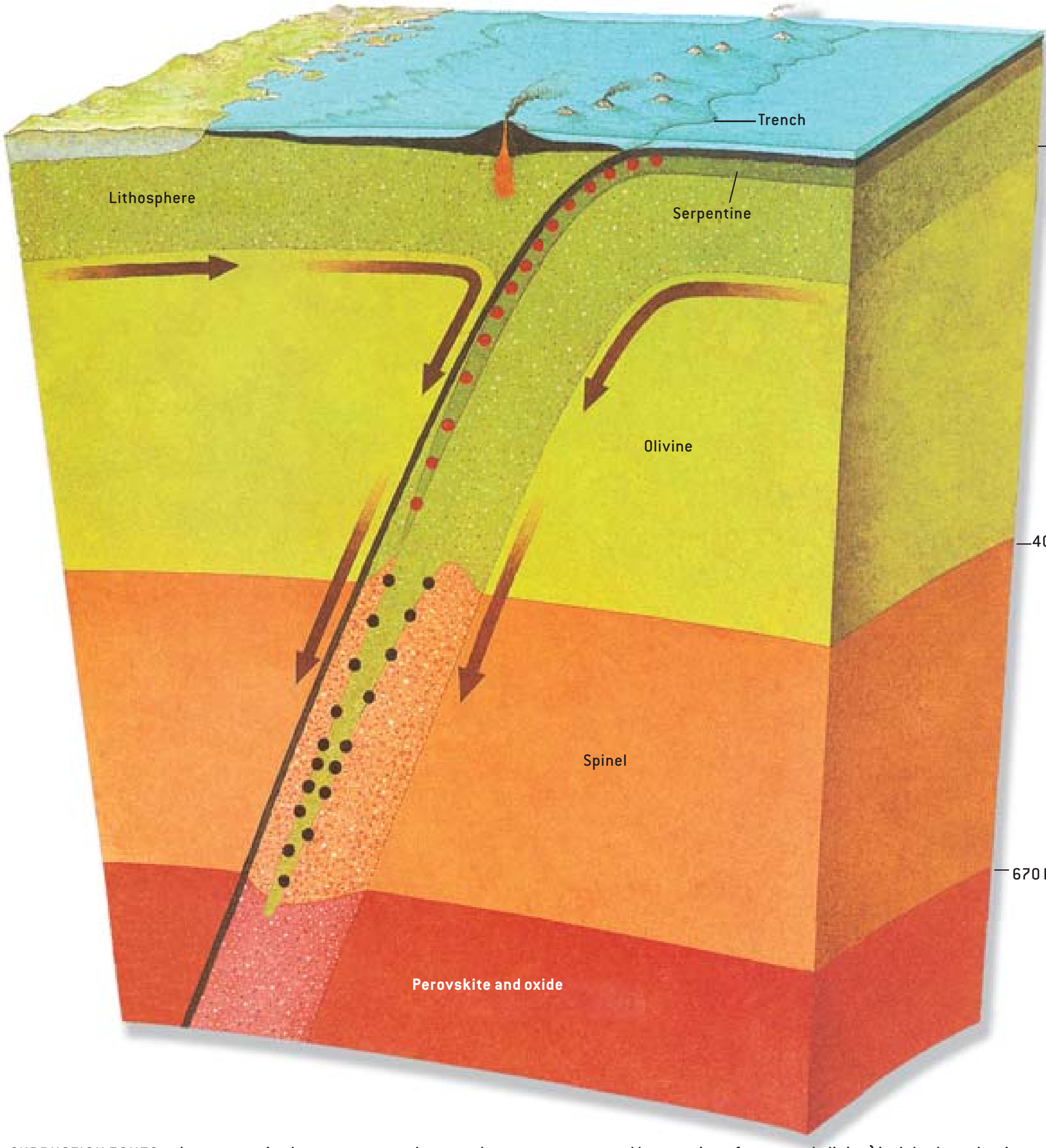
BECAUSE THE FREQUENCY of earthquakes steadily declines down to about 300 kilometers, most geophysicists believe that events originating between 70 and 300 kilometers below the surface (termed intermediate-focus earthquakes) are produced by a mechanism simply related to brittle fracture and frictional sliding. Deep-focus earthquakes (below 300 kilometers), however, follow an entirely different pattern and therefore probably stem from a separate mechanism. For more than six decades, the details of this mechanism remained elusive.

Years of study did provide intriguing information about subduction zones. Near Earth's surface, rocks contain minerals that exhibit a relatively loose packing of atoms. As the pressure on them increases at greater depths within the mantle, atomic reorganizations occur that yield new minerals having progressively greater density. The first such transformation occurs in most parts of the mantle at a depth of about 410 kilometers. In the reaction, olivine, the most abundant mineral of the upper mantle, becomes unstable and changes into a new phase having a spinel (cubic) structure that is 6 percent more dense than the original mineral. This shift causes an abrupt increase in seismic velocity at this depth. At 660 kilometers, the spinel form itself becomes unstable and decomposes into two new phases, which together are an additional 8 percent denser. The reaction induces another sharp rise in seismic velocity.

At lower temperatures, such as in a subducting slab, the spinel structure becomes stable at somewhat lower pressures than normal and remains so until subjected to slightly higher pressures than normal. Hence, in subduction zones the spinel stability field extends from a depth of about 300 kilometers to about 700 kilometers. This is exactly the region in which deep-focus earthquakes occur.

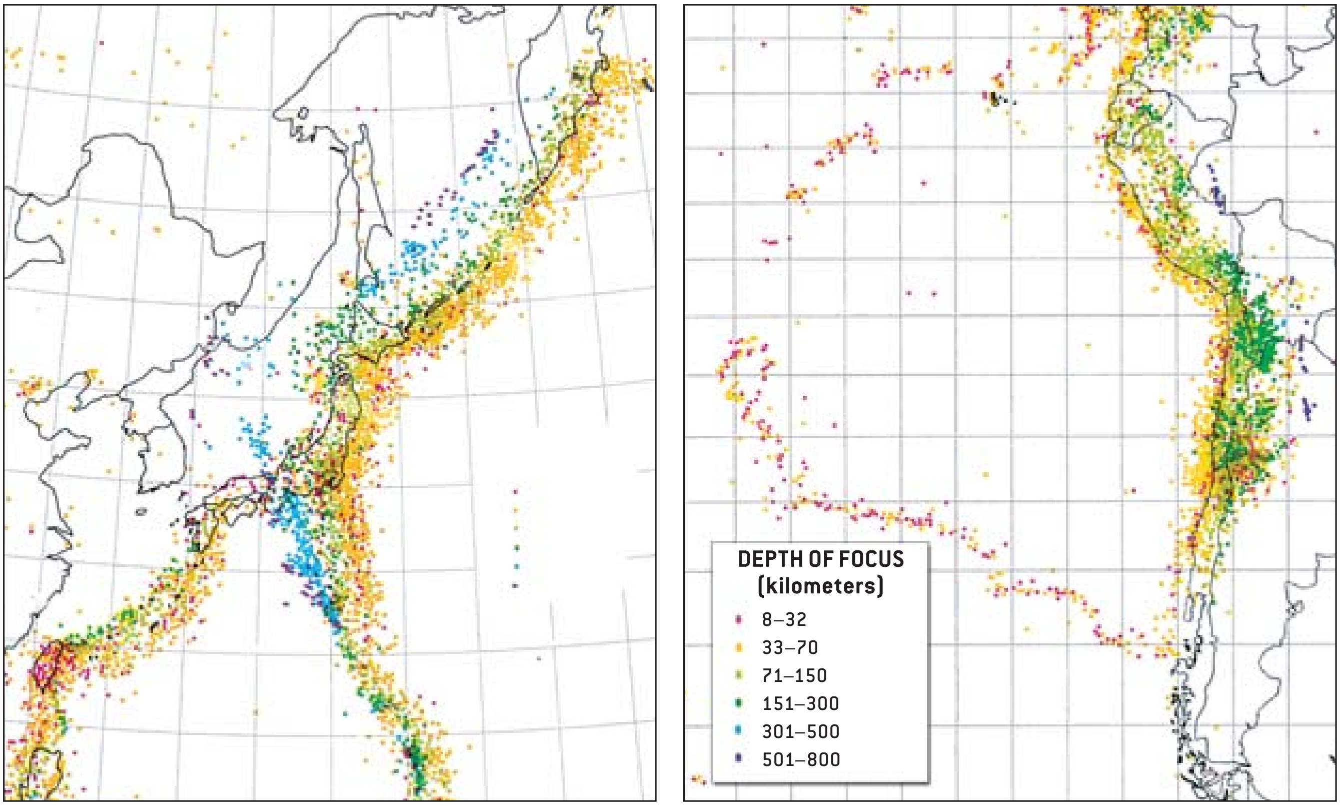
Because of this correlation, one of the recurring explanations over the years has been that the distribution of deep-focus earthquakes relates in some unknown way to these phase transformations. Most early suggestions focused on the fact that the reactions involve densification. Several researchers proposed that a sudden transformation of a significant volume of olivine to spinel would produce an implosion that could radiate the required seismic energy. Later studies refuted this hypothesis, however, showing that the geometric pattern of energy radiated from deep earthquakes is indistinguishable from that of shallow ones and indicating that movement takes place along a fault.

So what does cause deep earth-



SUBDUCTION ZONES, where tectonic plates come together, are the only places where intermediate- and deep-focus earthquakes occur. In contrast, shallow earthquakes happen throughout the world when the brittle rock in the lithosphere fractures and slides. In a cold subducted slab, different mechanisms promote deeper events. Intermediate-focus temblors (*red dots*) occur when serpentine (a hydrous mineral generated by reaction of water and olivine) is dehydrated as it descends into the mantle. Deep-focus quakes (*black dots*) result from the growth and spread of dense microstructures around the margins of the metastable olivine wedge (*green*) that extends below 400 kilometers. At 700 kilometers beneath the surface, any remaining olivine decomposes silently, and all earthquake activity stops.

ROBERTO OSTI



SEISMICITY mapped near Japan (*left*) and South America (*right*) from 1971 to 1986 shows that most earthquakes are shallow, originating at depths less than 70 kilometers beneath the earth's surface. Intermediate-focus earthquakes, those below 70 kilometers, and deep-focus events, those deeper than 300 kilometers, together make up only a third of all earthquakes.

U.S. GEOLOGICAL SURVEY

quakes, and why do the events correlate with the spinel stability field? Direct experimentation of any kind at the extraordinary pressures of Earth's deep interior has become possible only in the past few decades. In 1976 Roger Burns and Chien-Min Sung of the Massachusetts Institute of Technology showed that for temperatures and pressures expected in the cold core of a subduction zone, the transformation of olivine to spinel would probably be kinetically inhibited, even on a timescale of tens of millions of years. More recently, debate has arisen as to how cold a subduction zone would have to be to preserve such metastable olivine. Seismic evidence for the existence of metastable olivine, however, is now strong.

In the same year that Sung and Burns published their initial results, J. Rimas Vaisnys and Carol Pilbeam of Yale University suggested that a faulting instability might be possible during the transformation from olivine to spinel under certain conditions. In particular, they appealed to a thermal runaway (an exothermic reaction releases heat, which speeds the reaction rate, which generates more heat, and so on) and a marked decrease in crystal size, important characteristics that I will discuss further.

Despite such progress, in the late 1970s and early 1980s, a controversy arose concerning the exact mechanism by which olivine transforms to spinel. In addition to the silicate olivine of Earth's mantle, $(Mg,Fe)_2SiO_4$, the olivine-spinel transformation takes place in several other chemical systems, including germanate olivine, Mg_2GeO_4. The latter compound has the advantage that the larger germanium atom causes the transformation to occur at much lower pressures than it does in silicate olivine. Work in my laboratory using the germanate system agreed with the earlier observations of Sung and Burns—namely that the transformation occurred by the mechanism of nucleation and growth of spinel crystals on olivine grain boundaries. Studies elsewhere, though, suggested a different kind of mechanism, involving shearing of the crystal lattice. The differences between the various experiments caused me to suggest in 1984 that both mechanisms must be real and that stress probably determined which one would operate under a given set of conditions.

Looking for the Mechanism

IT WAS IMPORTANT to resolve the issue, because understanding the various aspects of mantle dynamics (including deep earthquakes) depends on knowing the mechanism responsible for this transformation. Thus, in 1985 Pamela C. Burnley (who was then a graduate student beginning her Ph.D. research) and I began investigating the effect of stress on the transformation. It was not possible at that time to perform deformation experiments and measure stress at the very high pressures under which this transformation takes place in mantle olivine. Therefore, Burnley (who is now at Georgia State University) and I continued to use magnesium germanate samples, because the pressure needed to induce the transformation was readily accessible in my experimental deformation machinery.

We prepared and deformed small

samples of a synthetic "rock" of this composition within the stability field of the spinel polymorph. The work confirmed that the level of stress determines the choice between the two mechanisms. At low temperatures, under conditions too cold for the reaction to run by nucleation and growth of new crystals, our specimens were very strong. They transformed only when high stress caused the crystal lattice to shear into thin lamellae of the denser phase. At high temperatures, however, the nucleation and growth mechanism ran quickly, and so the specimens were much weaker. In this case, the high stress that produced the shearing mechanism was never reached.

These results resolved the controversy over how olivine transforms into spinel. But the stresses required to produce the shearing mechanism are so high that only the nucleation and growth mechanism should operate in the earth. Moreover, we found no faulting instability associated with the shearing mechanism. Thus, it could be ruled out as a possible mechanism for deep earthquakes as well.

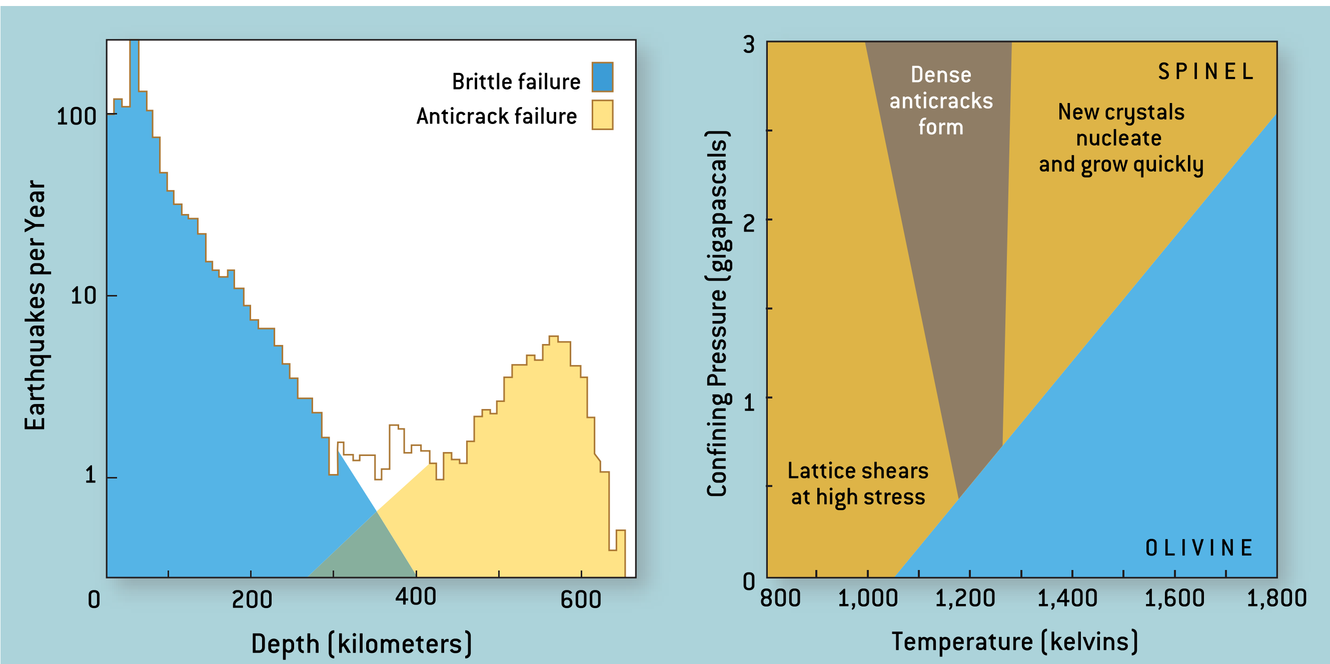
At the same time that Burnley was conducting these experiments, Stephen H. Kirby of the U.S. Geological Survey in Menlo Park, Calif., reported some anomalous results. They appeared in faulting studies of two minerals conducted near or above the pressures at which densification reactions might be expected. Although he found no direct evidence of such reactions, Kirby proposed that incipient transformation to the stable phases might have caused the faulting he observed. Like Vaisnys and Pilbeam 10 years earlier, he suggested that a faulting instability might operate in Earth's mantle during the transformation from olivine to spinel.

Although we had yet to witness this predicted instability, Burnley and I reasoned that if such an instability existed, it had to involve the nucleation and growth mechanism. Furthermore, the instability had to appear only in the narrow temperature interval between the two ranges tested during our earlier work. So, we deformed specimens under conditions for which nucleation of the spinel phase is just possible on the timescale of the experiment. Bingo! These specimens exhibited an abrupt drop in the amount of stress they could support and developed one or more spinel-lined faults cutting through them.

Detailed examinations revealed a unique set of microstructures within these faulted specimens. In the early stages of experiments conducted within the narrow faulting "window," microscopic packets of the high-density phase formed and grew on the olivine grain boundaries. These packets exhibited three critical characteristics: they looked like filled cracks; they were oriented perpendicular to the maximum compressive stress; and they contained extraordinarily small crystals of spinel (approximately 10^{-5} millimeter in diameter). The first two characteristics are tantalizingly similar to features that develop in brittle materials before they break. The third offered a potential answer as to how faults can form and slide at high pressures.

Forming a Theory

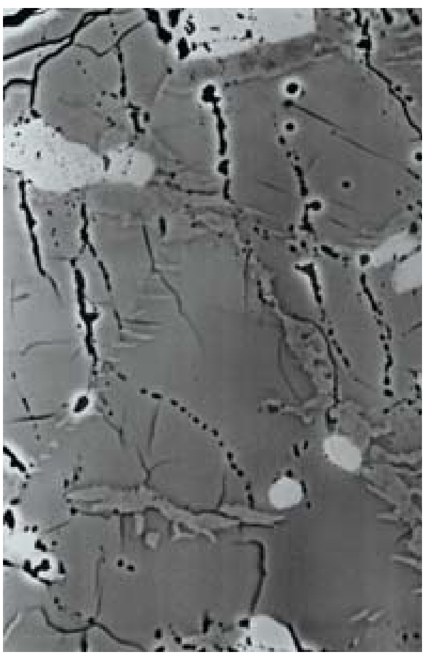
FROM THESE THREE characteristics, we formulated a theory of transformation-induced faulting that is analogous



LAURIE GRACE

FREQUENCY of earthquakes corresponds closely to the depths at which olivine undergoes phase transformations (*left*). A minimum number of events occur at roughly 400 kilometers, the depth at which olivine transforms into a denser spinel (or cubic) phase. No earthquakes occur below a depth of 700 kilometers, which is where spinel decomposes. Pressure and temperature govern these reactions (*right*). In germanate olivine that is subject to low pressures and high temperatures, olivine is stable, whereas at high pressures or low temperatures, the denser spinel is stable. Anticrack faulting occurs only in a narrow temperature "window."

to brittle shear fracture but differs fundamentally in its microphysics. In brittle shear fracture, as the stress rises, large numbers of microscopic tensile cracks open parallel to the maximum compressive stress. These features are referred to as Mode I cracks because the displacements across them are perpendicular to the plane of the crack. As loading continues, the number and density of Mode I microcracks increase rapidly until the material begins to lose its strength locally. At that time, the microcracks cooperatively organize to initiate shear fracture and the specimen fails in a fraction of a second. A "process zone" of tensile (Mode I) microcracks develops in front of the growing fault and leads it through the material. The important point here is that the fault is not a primary failure process; it must be prepared for and led by Mode I microcracks. Because pressure severely inhibits the expansion that takes place when tensile microcracks open, brittle failure cannot occur at depth.



MICROSCOPIC LENSES of dense spinel phase (*light gray*) in olivine weakened germanate olivine specimens. These "anticracks" weaken olivine and lead to failure. The rocks were deformed at temperatures at which spinel crystals nucleate sluggishly. At lower temperatures, high stress can induce a shearing transformation from olivine to the spinel phase. Likewise, at higher temperatures, spinel crystals nucleate easily and quickly.

In our high-pressure faulting experiments, we observed the growth of microscopic lenses of spinel in place of microcracks. The lenses are shaped very much like open tensile cracks, but they have the opposite orientation—they form perpendicular to the maximum compressive stress. The spinel phase is denser than olivine, hence the displacements of the lens boundaries move inward toward the plane of the lens. Therefore, the lenses are Mode I features like tensile cracks. But since the displacement of their boundaries is reversed, concentrations of compressive stresses develop at their tips rather than tensile stresses. It is the tensile stresses at the tips of opening cracks in brittle materials that cause them to orient themselves parallel to the maximum compressive stress; similarly, the compressive stresses at the tips of the lenses in our specimens cause them to orient themselves perpendicular to the maximum compressive stress.

Thus, in every way these features are the inverse of cracks—in a word, they are anticracks, a concept advanced in 1981 in a different context by Raymond Fletcher of Texas A&M University and David Pollard of Stanford University. Because of the remarkable similarities between the two Mode I features, we concluded that the microanticracks that precede failure in our experimental specimens must play the same role in high-pressure faulting as do microcracks in brittle fracture.

The third critical characteristic of our faulted specimens, the very fine-grained spinel in the fault zones, gave us insight as to how anticracks can provide a fundamental weakening step and why the process can occur at high pressure. Extremely fine-grained materials exhibit a remarkable flow property called superplasticity. Such materials flow by sliding on the grain boundaries between the crystals. This flow is somewhat like the deformation of a bag of sand but with the all-important difference that the grains of sand are rigid. Therefore, they must slide up and over one another. As gaps open up between sand grains, the dilation must work against the ambient pressure. Hence this process, like brittle failure, is severely inhibited by pressure. In contrast, grain-boundary sliding is a plastic process in which crystal defects called grain-boundary dislocations move. No dilatant movements occur (as in the granular flow of sand), so pressure has little inhibitory effect. Thus, we postulated that the fine-grained spinel within the anticracks is much weaker than the host olivine and has this "superplastic" flow capacity.

From these observations we formulated the following hypothesis: during loading, under conditions for which the spinel phase grows with difficulty, olivine transforms to spinel primarily as new crystals that form by repeated nucleation adjacent to one another at concentrations of stress. In a nonhydrostatic stress field, the developing packets of spinel tend to grow perpendicular to maximum compressive stress. This preference leads to their lens-shaped morphology and general orientation. These Mode I microanticracks initially develop randomly throughout the specimens. But because the fine-grained spinel aggregates within the microanticracks are much weaker than the large olivine crystals, once enough of them have formed, the specimen loses its strength locally.

At this critical stage, large stress concentrations develop around the region of incipient failure, and the growth of an-

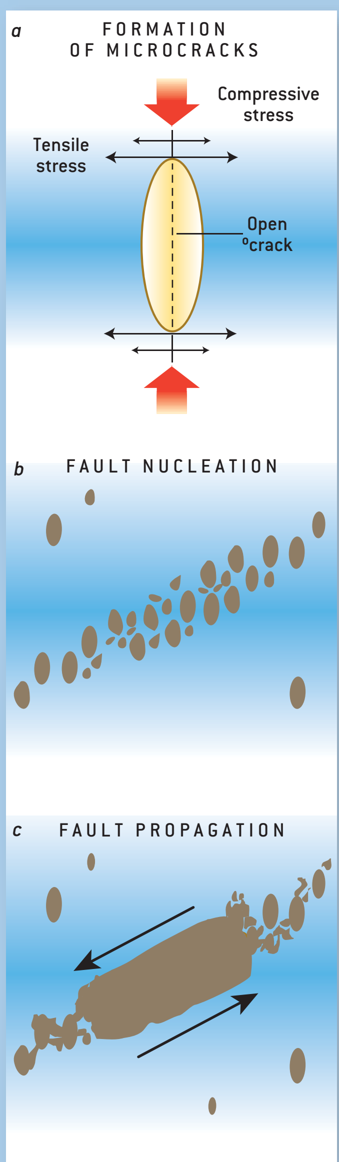
THE AUTHOR

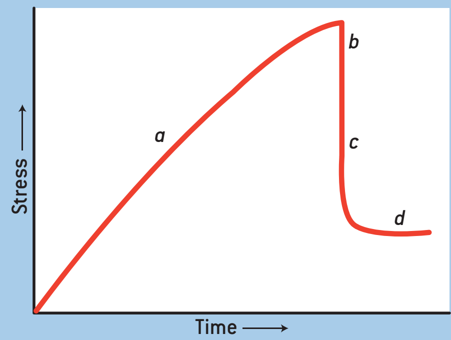
HARRY W. GREEN II is professor at the University of California, Riverside, where he is a member of the Institute of Geophysics and Planetary Physics. In 1968 he received a Ph.D. in geology and geophysics from the University of California, Los Angeles. Currently he investigates the rheology of geologic materials, how stress affects polymorphic phase transformations, and the return of subducted rocks to the surface from great depth. He is a fellow of the American Geophysical Union, the Mineralogical Society of America and the American Association for the Advancement of Science. He has published more than 100 papers and offers counsel to the National Science Foundation, the Department of Energy and other such agencies.

HARRY W. GREEN II

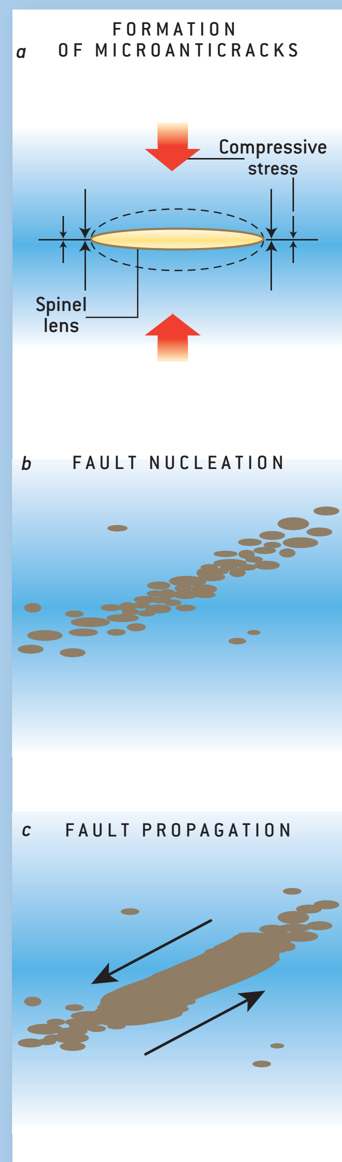
BRITTLE VERSUS ANTICRACK FAILURE

BRITTLE FAILURE





ANTICRACK FAILURE



Brittle failure and anticrack failure, the mechanisms that account for shallow and deep earthquakes, respectively, share many characteristics. Both processes involve the development of microscopic features (*a*) that cooperatively form a fault and allow for movement on it.

As brittle rocks experience mounting stress, microcracks open parallel to the direction from which they are compressed. In rocks found more than 300 kilometers deep, pressure prohibits this dilation. Instead microanticracks (lenses filled with fine-grained, dense spinel-phase olivine) form perpendicular to the direction from which these rocks are compressed.

In each mechanism, at some critical point in time, these microfeatures link up and create a fault (*b*). Movements along the fault then relieve stress (*c*). In brittle failure, the fault is an open fracture containing broken rock fragments. In anticrack failure, the fault contains fine-grained spinel-phase olivine. This material is superplastic—that is, the crystals can readily move past one another, enabling the fault to slide. Because the fault need not dilate to slip in this way, pressure does not restrict the process.

Brittle fault zones contain angular crystals showing a fractal size distribution. In anticrack fault zones, rounded olivine fragments are embedded in extremely fine grained spinel-phase olivine (*d*).

d FAULT ZONE

d FAULT ZONE

IAN WORPOLE; HARRY W. GREEN II (*bottom, left and right*) ticracks accelerates. Preexisting microanticracks then link up and empty their superplastic contents into the developing fault zone, providing a lubricant along which the fault can slide. The process continues ahead of the tip of the growing fault zone and thereby provides the superplastic material needed to lubricate the fault. The anticracks must grow very rapidly to produce this faulting. We postulated that the speed of their growth resulted from a thermal feedback mechanism: the nucleation of spinel in the anticracks releases heat that locally increases the temperature, which increases the nucleation rate, which raises the temperature further, and so forth, leading to catastrophic failure.

Putting Theory to the Test

BURNLEY AND I published the essence of this model in Nature magazine in October 1989, and much of the time in my laboratory since then has been spent testing various aspects of the theory. Happily, it has survived all our scrutiny thus far. In one very important test, we investigated whether energy is radiated elastically during anticrack faulting. Obviously, if anticrack faulting is "silent," it cannot be responsible for earthquakes, because the shaking we experience is a result of the arrival of "noise" emitted during the failure process. Because our specimens were small and located deep within the deformation apparatus (which itself produces general background noise), we could not hear the sound emitted during the faulting process.

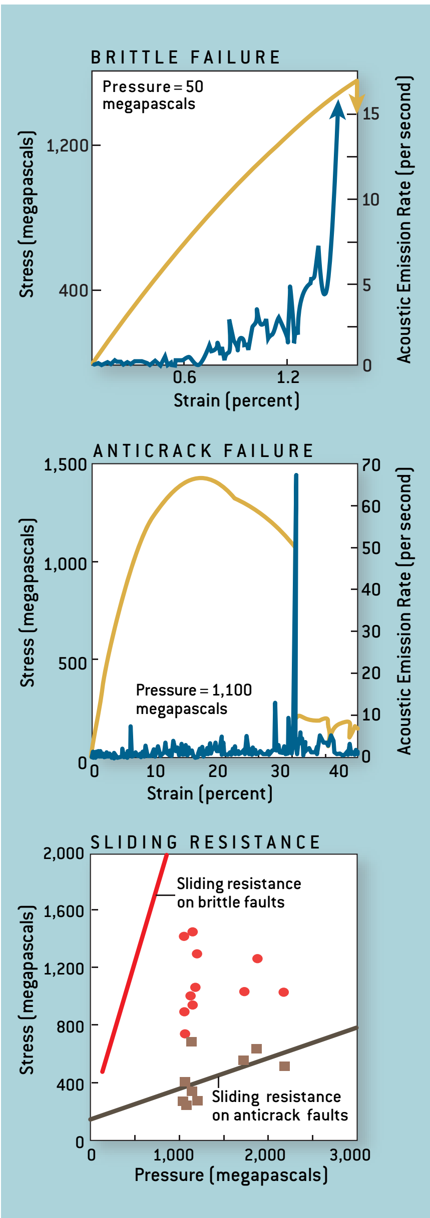
Therefore, I established a collaboration with Christopher H. Scholz of the Lamont-Doherty Earth Observatory of Columbia University, who investigates brittle fracture. Scholz attaches sensitive piezoelectric transducers to his apparatus to "listen" to the acoustic emissions that precede and accompany brittle failure. We modified one of my high-pressure deformation apparatuses to reduce noise and—working with Tracy N. Tingle (now deceased) and Thomas E. Young from my lab and Theodore A. Kozynski from Scholz's lab—successfully detected acoustic emissions from samples of Mg_2GeO_4 during failure.

Tingle and I investigated the flow strength of Mg_2GeO_4 spinel when the crystals are comparable in size to the olivine crystals of the starting material. We then compared that strength with the resistance against sliding present on anticrack-induced faults. Whereas this resistance is much less than the flow strength of the olivine specimens before failure, the flow strength of coarse-grained spinel is twice as great as that of olivine. As a result, one cannot explain the weakness of the fault zones in our specimens by simply replacing olivine with spinel; the flow mechanism must also change. The only known mechanism that can provide such weakening is superplastic flow, consistent with our original speculation.

These tests established beyond doubt that anticrack faulting was a new failure mechanism distinct from brittle failure. Nevertheless, they had one major flaw: we conducted all of these experiments on germanate olivine, not the silicate olivine found in the mantle. Of course, as mentioned above, none of this work could have been done on silicate olivine; at that time it was impossible to measure stresses at the high pressures needed to reach the spinel stability field in the silicate system.

David Walker, also at Lamont-Doherty, then suggested that we attempt crude deformation experiments on mantle olivine in his multianvil apparatus, a machine that can attain the requisite pressures. Such a device had never before been used for deforming mineral specimens, but we decided to give it a try. Our philosophy was that if anticrack faulting truly gives rise to deep-focus earthquakes, it must operate in real olivine. The microstructures we observed in the germanate specimens could guide us to uncover the conditions under which faulting would develop in the silicate. The approach worked beyond our wildest dreams; after only four trials, we produced faulting and characteristic anticrack microstructures at a pressure of 14 gigapascals.

Despite the attractive properties of the anticrack faulting mechanism, it can operate only if olivine is carried deep into the upper mantle, where the spinel crystal structure is stable. In particular,



ACOUSTIC EMISSIONS (*blue*) occur when a fault suddenly slips, radiating energy that relieves stress (*yellow*). Brittle failure, the mechanism responsible for shallow earthquakes, emits noise before and during movement on a fault (*top*). In contrast, anticrack failure, the mechanism behind deep earthquakes, emits acoustic energy only when faulting takes place (*middle*). At higher pressures, brittle faulting can occur only under much greater stress (*bottom*). For this reason, brittle failure cannot explain deep-focus earthquakes. Pressure does not likewise inhibit anticrack failure.

LAURIE GRACE

this mechanism cannot account for earthquakes shallower than approximately 300 kilometers, where olivine is still stable. Normal brittle fracture, though, cannot explain earthquakes deeper than 70 kilometers. What transpires in between these depth regions? Other experiments have neatly provided the explanation for intermediate-focus earthquakes.

Barry Raleigh, now at the University of Hawaii at Manoa, and Mervyn Paterson of the Australian National University in Canberra demonstrated in the 1960s that when the hydrous mineral serpentine breaks down to olivine plus water under stress at somewhat elevated pressure, it enables brittle shear failure under conditions where the rock otherwise would flow. We can understand this process in terms of the anatomy of brittle fracture: The pressure of the water produced by dehydration pushes open microcracks against the high applied pressure, thereby allowing for brittle failure. Little additional experimental work was produced on this system for the following 40 years until Haemyeong Jung, Larissa Dobrzhinetskaya and I showed in 2004 that this faulting mechanism is viable to at least six gigapascals of pressure, despite the fact that the volume change during serpentine dehydration changes from expansion to contraction.

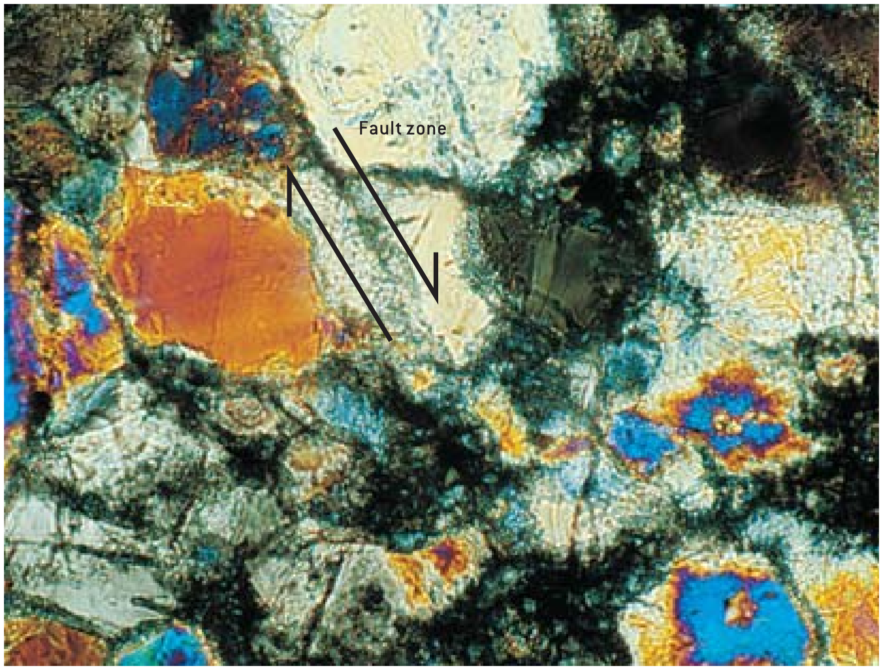
We know from a variety of geophysical and geological observations that olivine in the uppermost mantle (just below the oceanic crust) becomes partially hydrated as it journeys from an ocean ridge to an ocean trench. Thus, shallow regions in the lithosphere contain the hydrous minerals that enable this mechanism to work. The declining frequency of earthquakes in subduction zones down to 300 kilometers most probably represents the progressive exhaustion of this mechanism as the oceanic lithosphere gradually warms up and dehydrates, heated by the surrounding mantle. At about 300 kilometers, anticrack faulting becomes possible, causing an increase in earthquakes there.

The anticrack faulting mechanism provides an explanation for how and why earthquakes extend to great depth. But can this mechanism also explain why they suddenly stop? As mentioned above, the decomposition of spinel into two denser phases occurs at approximately 700 kilometers deep in subduction zones. This decomposition reaction is endothermic (it requires the addition of heat to proceed). In contrast, the transformation from olivine to spinel is exothermic (heat is released during the reaction). If we were correct in our original assumption that a thermal runaway must occur to introduce a faulting instability, then an endothermic reaction should be incapable of producing such an instability.

To test this possibility, in 1996 Yi Zhou and I published a set of experiments on $CdTiO_3$, a composition that undergoes an endothermic densification transformation. Deformation of the low-pressure phase under conditions for which the high-pressure phase is stable proceeded uneventfully; we observed neither anticracks nor faulting. If olivine should be carried all the way to 700 kilometers or more, however, its direct breakdown in the lower mantle would be exothermic. Therefore, Gayle Gleason (now at the State University of New York College at Cortland) and I performed a series of experiments on the mineral jadeite ($NaAlSi_3O_8$) that breaks down to two dense phases in a strongly exothermic reaction. Once again, no faulting was produced. These two sets of experiments demonstrated that anticrack faulting requires both an exothermic reaction *and* a single product phase. Thus, earthquakes triggered by this mechanism are not possible in the lower mantle.

In summary, the depth distribution of earthquakes and the experimental results lead naturally to the following model. Normal brittle fracture and frictional sliding accounts for shallow earthquakes. Because pressure inhibits this mechanism, earthquakes in most parts of the world cease by 20 to 30 kilometers below Earth's surface. In subduction zones, however, partially hydrated oceanic crust and mantle sink downward and become slowly heated. The water-bearing minerals begin to dehydrate and, in the process, make fluid-assisted faulting possible. The exponential decrease in earthquake frequency down to 300 kilometers reflects the progressive heating and dewatering of the subducting slab.

The interior of the slab remains suf-



FAULT ZONE, produced by deforming a sample of mantle olivine, extends from the top left corner to the bottom right corner of the micrograph. An offset olivine crystal (*white*) appears on this fault in the top half of the image. The anticracks (*yellow lenses in blue crystal*) that generate the fault zone form perpendicular to the direction from which the sample is compressed. They grow throughout the material parallel to this direction.

HARRY W. GREEN II

ficiently cold so that the olivine of the subducting mantle cannot transform to the spinel phase when it leaves the olivine stability field at approximately 300 kilometers deep. At the margins of this cold interior region, the temperature slowly increases. The metastable olivine heats to the critical temperature at which anticrack faulting occurs. In the coldest subduction zones, the wedge of metastable olivine extends down about 700 kilometers, generating earthquakes all the way. Earthquakes, however, stop there whether or not the slab enters the lower mantle; olivine decomposition into the two very dense phases of the lower mantle would be silent.

The seismic velocity of the cold interior of subducting plates should be significantly slower if metastable olivine is present than if the reaction has already run to produce the denser polymorphs. In 1992 Takashi Iidaka and Deisuke Suetsugu of the University of Tokyo modeled both possibilities for the descending slab beneath Japan and found the telltale slow velocity of the metastable olivine wedge. This result has remained controversial, but recent results by Wang-Ping Chen and Michael Brudzinski of the University of Illinois have shown that a peculiar extended subhorizontal group of deep earthquakes west of the Tonga subduction zone in the southwest Pacific Ocean occur in rocks whose seismic velocity is much too slow to be normal for rocks at that depth; they conclude that the only interpretation consistent with all the observations is that these earthquakes are occurring in a fossil subducted slab that is floating at 400 to 600 kilometers, its buoyancy and earthquake activity caused by extensive retention of metastable olivine.

If, as we propose, a critical temperature controls the anticrack faulting instability, the faulting in descending slabs will be concentrated at the interface between the metastable olivine wedge and the surrounding, already transformed carapace. If sufficient stress exists on both margins of the wedge, double zones of earthquakes could develop. Two sets of seismologists, one led by Douglas Wiens of Washington University and the other by Iidaka, discovered such double zones in 1994. Iidaka's team found the double zone in the slab in which they previously reported a metastable olivine wedge.

Additional indirect seismic evidence of the mechanism of deep earthquakes comes from remotely triggered deep earthquakes, also beneath Fiji, reported in 2003 by Rigobert Tibi of Washington University, Wiens and Hiroshi Inoue of the National Research Institute for Earth Science and Disaster Prevention in Japan. They showed that the seismic waves from a large earthquake, magnitude 7.6, triggered two other earthquakes (one of magnitude 7.7) in an area 300 kilometers away and 60 kilometers deeper within seven minutes of the first earthquake in a region of no previous recorded earthquake activity. By the time the earthquake waves of the first large earthquake had traveled 300 kilometers, the stresses and strains associated with those waves would have become very small, yet they were sufficient to set the earthquake process in motion at that great distance—implying that the regions where the triggered earthquakes occurred were ripe for failure and needed only the smallest of assistance to be set off; but even so, a few minutes were needed for local organization of the process. These same authors then showed equally compelling evidence that another pair of large earthquakes in 1986, separated in time by 26 minutes and in distance by hundreds of kilometers, must also have been a triggering/triggered pair.

In a discussion of this paper, I pointed out that of the three mechanisms previously proposed to explain deep earthquakes (dehydration embrittlement, transformation-induced faulting and runaway shear heating), the first two are known from experiments to be potentially activated within a few minutes and hence cannot be ruled out by these observations. Even under optimal conditions, however, runaway shear heating would require a minimum of several years and more probably thousands of years to develop, making it highly unlikely that two locations in a region never before found to generate earthquakes would have just happened to be on the verge of failure.

In conclusion, the laboratory results explain how earthquakes can be initiated at very high pressures. The composite model advanced here, in which intermediate-depth earthquakes are triggered by fluid-assisted faulting and deep events by anticrack faulting, is highly consistent with our current understanding of subduction zones. But any earthquake could develop a component of shear heating during its propagation, thereby perhaps contributing significantly to the very large, very deep earthquakes, as suggested by Hiroo Kanamori, Don Anderson and Thomas Heaton of the California Institute of Technology for the great Bolivian earthquake of 1994. Modern seismological techniques are now illuminating this problem by providing methods to detect the presence of metastable olivine or other mineralogical anomalies that may be involved in generation of deep earthquakes. Thus, questions remain, but the paradox behind deep earthquakes has been resolved. SA

MORE TO EXPLORE

A New Self-Organizing Mechanism for Deep-Focus Earthquakes. Harry W. Green II and Pamela C. Burnley in *Nature,* Vol. 341, pages 733–737; October 26, 1989.

Anticrack-Associated Faulting at Very High Pressure in Natural Olivine. Harry W. Green II, Thomas E. Young, David Walker and Christopher H. Scholz in *Nature,* Vol. 348, pages 720–722; December 20–27, 1990.

The Mechanics of Deep Earthquakes. Harry W. Green II and Heidi Houston in *Annual Reviews of Earth and Planetary Sciences,* Vol. 23, pages 169–213; 1995.

Evidence for a Large-Scale Remnant of Subducted Lithosphere Beneath Fiji. Wang-Ping Chen and Michael R. Brudzinski in *Science,* Vol. 292, pages 2475–2479; June 29, 2001.

Instability of Deformation. Harry W. Green II and Chris Marone in *Plastic Deformation of Minerals and Rocks.* Edited by Hans-Rudolph Wenk and Shun-Ichiro Karato. *Reviews in Mineralogy and Geochemistry,* Vol. 51, 2002.