ECOSCIENCE: POPULATION, RESOURCES, ENVIRONMENT

PAUL R. EHRLICH
STANFORD UNIVERSITY
ANNE H. EHRLICH
STANFORD UNIVERSITY
JOHN P. HOLDREN
UNIVERSITY OF CALIFORNIA, BERKELEY

W. H. FREEMAN AND COMPANY
San Francisco

1977

All the rivers run into the sea; yet the sea is not full.

– Ecclesiastes

CHAPTER 2

The Physical World

The *biosphere* is that part of Earth where life exists. In vertical dimension it extends from the deepest trenches in the ocean floor, more than 11,000 meters (36,000 feet)[1] below sea level, to at least 10,000 meters (m) above sea level, where spores (reproductive cells) of bacteria and fungi can be found floating free in the atmosphere. By far most living things – most of which depend directly or indirectly on the capture of solar energy by photosynthesis in plants and certain bacteria – exist in the narrower region extending from the limit of penetration of sunlight in the clearest oceans, less than 200 meters from the surface, to the highest value of the permanent snow line in tropical and subtropical mountain ranges – about 6000 meters, or 20,000 feet. (Everest, the highest mountain, rises almost 8900 meters above sea level.) By any definition, the biosphere is as a mere film in thickness compared to the size of the ball of rock on which it sits – about like the skin of an apple, in fact. The radius of Earth is about 6370 kilometers (km), or 4000 miles (mi).

Of course, conditions within the thin envelope of the biosphere are influenced by physical processes taking place far outside it: by the energy emitted by the sun, 150

---

[1] Throughout this book physical dimensions are given in metric units, sometimes accompanied by the English equivalent to ease the transition for readers not completely accustomed to the metric system. For more precise conversion factors, see the tables inside the covers of the book.

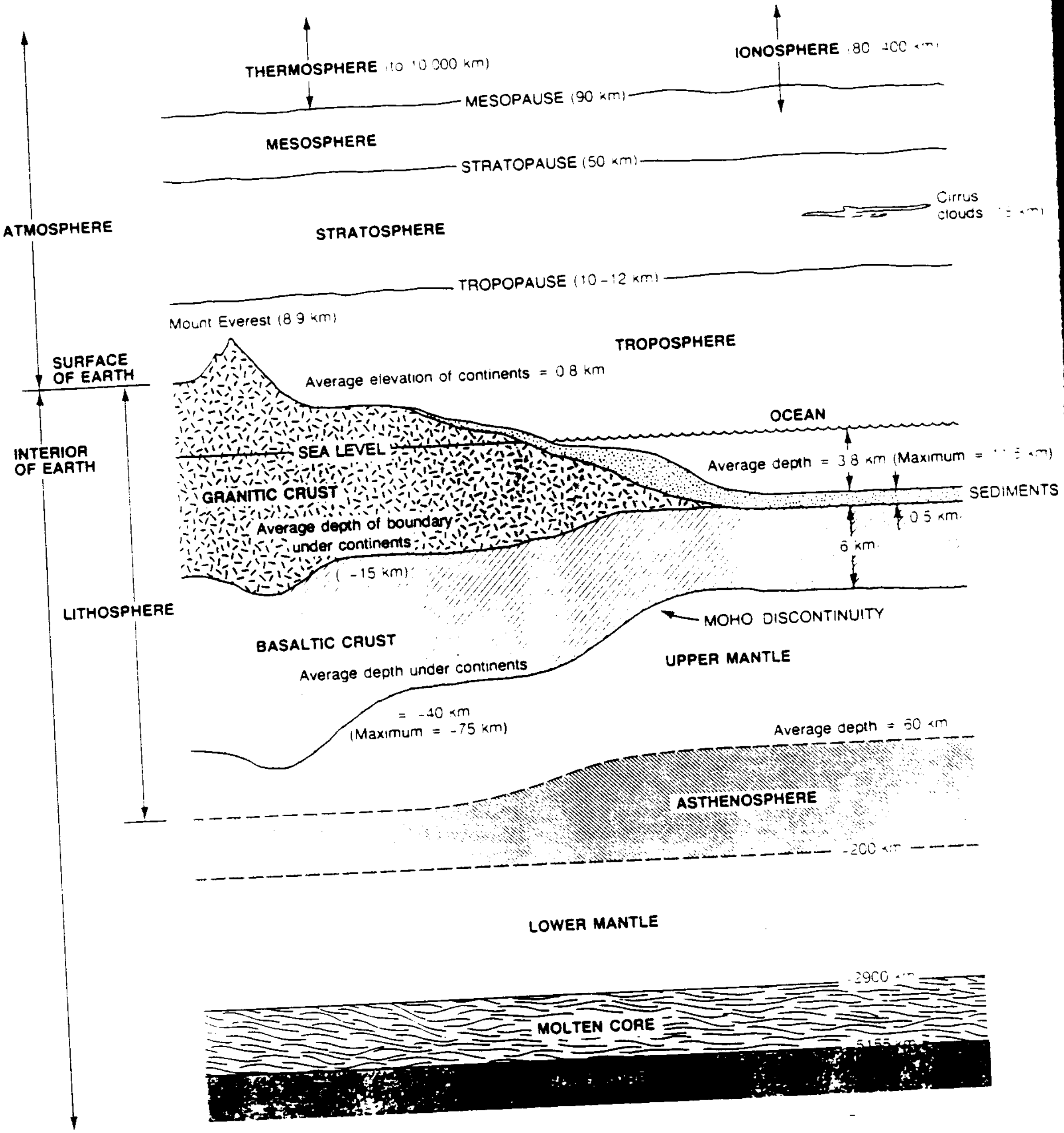


FIGURE 2-1

Vertical structure of the physical world. (Scales are greatly distorted.)

million kilometers away; by the tides originating in the relative motion and positions of Earth, sun, and moon (the distance from Earth to the moon averages about 380,000 km); by the presence of gases 20 to 400 km high in the atmosphere that screen out harmful components of incoming solar energy; by the constitution and structure of Earth's crust (to 40 km deep), which govern the availability of mineral nutrients at the surface and of metallic ores accessible to industrial civilization; by the behavior of the solid but plastically flowing[2] mantle (to 2900 km deep) on which the crustal plates "float" and move laterally; and by the motion of Earth's molten core (2900 to 5200 km deep), which produces the magnetic field that protects the planet's surface from bombardment by energetic, electrically charged particles from space. The vertical structure of Earth's atmosphere, surface, and interior is illustrated in Figure 2-1. The terminology used there for the various vertical divisions is explained in the following text, where the character of the atmosphere and Earth's interior are taken up in more detail.

In horizontal extent the biosphere covers the globe, although in the hottest deserts and coldest polar regions—as at the highest elevations—usually only dormant spores can be found. Earth's total surface area amounts to 510 million square kilometers (about 197 million mi²), of which 71 percent is ocean and 29 percent land (see Table 2-1). The mass of all living organisms on Earth amounts to about 5 trillion metric tons,[3] three-fourths of which consists of water. Under the reasonable assumption that living matter has about the same density as water [1 gram per cubic centimeter (1g/cm³)], this would mean that the living part of the biosphere was equivalent to a layer of material only 1 centimeter thick, covering the globe. (The range concealed in this average is from 0.0002 g of living material for each square centimeter of surface in the open ocean to 15 g or more for each square centimeter of surface in a tropical forest.)

---

[2] *Plastic flow* in a solid refers to continuous deformation in any direction without rupture.

[3] A metric ton (MT) equals 1000 kilograms (kg), or 2205 pounds. One trillion (in scientific notation, 10^{12}) equals 1000 billion, or 1,000,000 million.

TABLE 2-1
Surface Areas of the Globe

	Percentage of category	*Area (million km²)*
TOTAL EARTH SURFACE		510
OCEANS		361
Pacific Ocean*	46	
Atlantic Ocean*	23	
Indian Ocean*	18	
Arctic Ocean*	4	
Mean extent of sea ice (Arctic, South Atlantic, Pacific, and Indian)	7	
LAND		149
Eurasia	36	
Africa	20	
North and Central America	16	
South America	12	
Antarctica	10	
Oceania	6	
Ice-covered land	10	
Terrain more than 3000 meters high	5	
Lakes and rivers	1	

*Mean extent of sea ice has been subtracted.
Source: Strahler and Strahler, *Environmental geoscience.*

TABLE 2-2
Masses of Constituents of the Physical World

Constituent	*Mass (trillion MT)*
Living organisms (including water content)	8
Liquid fresh water, on surface	126
Atmosphere	5,140
Ice	30,000
Salts dissolved in oceans	49,000
Oceans	1,420,000
Earth's crust (average depth, 17 km)	24,000,000
Earth (total)	6,000,000,000

The relative masses of various constituents of the physical world are shown in Table 2-2.

To study the processes that operate in any subdivision of the physical world—atmosphere, biosphere, Earth's crust—one must know something about energy: what it is, how it behaves, how it is measured. For whenever and wherever anything is happening, energy in some form is involved; it is in many respects the basic currency of the physical world. An introduction to energy and the related concepts of work and power, along with the units in which these quantities can be measured, is provided in Box 2-1. Some feeling for how much energy is stored in and flows between various parts of the physical world is conveyed in Table 2-3.

EARTH'S SOLID SURFACE AND BELOW

The outermost layer of Earth's solid surface is called the ***crust.*** It ranges in thickness from about 6 kilometers beneath the ocean floor to as much as 75 kilometers below the largest mountain ranges. In essence, the crust floats on the denser ***mantle*** beneath it. (As is elaborated below, the crust and mantle are differentiated by the different compositions and densities of the rock they comprise.) As with icebergs on the sea, the more crust extends above the surface (as in a mountain range), the more bulk is hidden below.[4] This situation is made possible by the existence of a soft, yielding layer called the ***asthenosphere*** in the middle of the underlying mantle. This layer's strength is low because the rock is near its melting point. The combination of the crust and the hard upper layer of the mantle is called the ***lithosphere,*** a term sometimes also employed in a more general sense to mean the entire solid part of Earth. Below the mantle, between the depths of 2900 and 5200 kilometers, lies Earth's ***molten outer core.*** This core consists largely of liquid iron (with some nickel) at a temperature of perhaps 2500° C; its properties and motion produce Earth's magnetic field.

TABLE 2-3
Energy Flow and Storage in the Physical World

	Energy or power
STORAGE	*Trillion MJ*
Energy released in a large volcanic eruption	100
Chemical energy stored in all living organisms	30,000
Energy released in a large earthquake	100,000
Chemical energy stored in dead organic matter	100,000
Heat stored in atmosphere	1,300,000
Kinetic energy of Earth's rotation on its axis	250,000,000,000
FLOWS	*Million Mw*
Tides	3
Heat flow from Earth's interior	32
Conversion of sunlight to chemical energy in photosynthesis	100
Conversion of sunlight to energy of motion of atmosphere	1,000
Sunlight striking top of atmosphere	172,000

---

[4] For more detailed treatment of this point and others in this section, see F. Press and R. Siever, *Earth;* and A. N. Strahler and A. H. Strahler, *Environmental geoscience.*

BOX 2-1 Work, Energy, and Power: Definitions, Disguises, and Units

Work is the application of a force through a distance. *Energy* is stored work. *Power* is the rate of flow of energy, or the rate at which work is done. All these concepts are more easily understood with the help of examples and some elaboration.

Work – force multiplied by distance – is done when a weight is lifted against the force of gravity (as with water carried upward in the atmosphere in the course of the hydrologic cycle), when mass is accelerated against the resistance of inertia (as with waves whipped up on the ocean by the wind) or when a body is pushed or pulled through a resisting medium (as with an aircraft moving through the atmosphere or a plow cutting through a field). The presence of distance in the concept of work means that work is done only if there is motion – if you push on a stalled car and it doesn't budge, there is a force, but there is no work because there is no motion.

The foregoing are examples of *mechanical work* – work involving the bulk (or *macroscopic*) motion of agglomerations of molecules. There are also various forms of *microscopic* work, such as *chemical work* and *electrical work*, which involve forces and motions on the scale of individual molecules, atoms, and electrons. To heat a substance is to do a form of microscopic work in which the individual molecules of the substance are made to move more rapidly about in all directions, without any bulk motion taking place. The demonstration that all these different manifestations of work are fundamentally the same can be found in treatises on physics and chemistry.*

If work has many guises, so must energy, which is only *stored* work. Work stored as the motion of a macroscopic object (for example, a speeding automobile or the Earth spinning on its axis) is called *mechanical energy* or *kinetic energy*. The latter term may be applied as well to the energy of motion of microscopic objects (such as, molecules, electrons). Work stored as the *disordered* motion of molecules – that is, rotation, vibration, and random linear motion not associated with bulk motion of the substance – is called *thermal energy* or *sensible heat* or (more commonly) just *heat*. Note that temperature and heat are not the same. Temperature is a measure of the intensity of the disordered motion of a typical molecule in a substance; the heat in a substance is the sum of the energies stored in the disordered motion of all its molecules. (The relation between temperature and energy is developed further in Box 2-3.)

Kinetic energy means something is happening; that is, the work is stored as motion. *Potential* or *latent* energy means something is "waiting" to happen. That is, the work is stored in the position or structure of objects that are subject to a force and a restraint; the force provides the potential for converting position or structure into kinetic energy, and the restraint is what keeps this from happening (at least temporarily). Each kind of potential energy is associated with a specific kind of force. Gravitational potential energy (an avalanche waiting to fall) and electrical potential energy (oppositely charged clouds waiting for a lightning stroke to surge between them) are associated with forces that can act between objects at large distances. Chemical potential energy (gasoline waiting to be burned, carbohydrate waiting to be metabolized) is associated with the forces that hold atoms together in molecules – that is, with chemical bonds. Nuclear potential energy is due to the forces that hold protons and neutrons together in the nucleus – the so-called strong force. Latent heat of vaporization (water vapor waiting to condense into liquid, whereupon the latent heat will be converted to sensible heat) and latent heat of fusion (liquid waiting to freeze into a solid, with the same result) are associated with the electrical forces between molecules in liquids and solids. The idea that potential energy is something "waiting" to happen needs only to be tempered by recognition that sometimes it can be a long wait – the chemical potential energy in a piece of coal buried in Earth's crust, for example, may already have waited a hundred million years.

Electromagnetic radiation is a form of energy that does not fall neatly into any of the categories we have mentioned so far. It is characterized not

(*Continued*)

---

*See, for example, R. Feynmann, R. Leighton, and M. Sands, *The Feynmann lectures on physics*, Addison-Wesley, Reading, Mass., 1965.

BOX 2-1 (*Continued*)

in terms of the motion or position or structure of objects but in terms of the motion of electric and magnetic forces.[**] Light (visible electromagnetic radiation), radio waves, thermal (infrared) radiation, and X-rays are all closely related varieties of this particular form of energy. (See also Box 2-4.)

Albert Einstein theorized, and many experiments have subsequently verified, that any change in the energy associated with an object (regardless of the form of the energy) is accompanied by a corresponding change in mass. In this sense, mass and energy are equivalent and interchangeable, the formal expression of equivalence being Einstein's famous formula $E = mc^2$. (Here E denotes energy, m mass, and c the speed of light.) Because a small amount of mass is equivalent to a very large quantity of energy, a change in mass is only detectable when the change in energy is very large – as, for example, in nuclear explosions.

Different professions use a bewildering array of units for counting work and energy. The metric system is prevailing, but so gradually that the literature of energy and environmental sciences will be littered for years to come by a needless profusion of archaic units. Since work has the dimensions of force times distance, and all energy can be thought of as stored work, it should be apparent that a single unit will suffice for all forms of energy and work. The most logical one is the *joule* (J), which is exactly the amount of work done in exerting the basic metric unit of force, 1 *newton* (N), over the basic metric unit of distance, 1 meter.[†] We shall use the joule and its multiples, the kilojoule (1000 J, or kJ) and the megajoule (1,000,000 J, or MJ), throughout this book.

Our only exception to the use of the joule is a concession to the enormous inertia of custom in the field of nutrition, where we reluctantly employ the *kilocalorie* (kcal). A kilocalorie is approximately the amount of thermal energy needed to raise the temperature of 1 kg of water by 1 degree Celsius (1° C);[††] this unit is often confusingly written as "calorie" in discussions of nutrition. Running the bodily machinery of an average adult human being uses about 2500 kcal – about 10,000 kJ – per day.)

Besides the erg (1 ten-millionth of a joule) and the calorie (1 thousandth of a kcal), the unit of energy most likely to be encountered by the reader elsewhere is the British thermal unit (Btu), which is approximately the amount of thermal energy needed to raise 1 pound of water by 1 degree Fahrenheit (1° F). A Btu is roughly a kilojoule.

In many applications one must consider not only *amounts* of energy but the *rate* at which energy flows or is used. The rate of energy flow or use is *power*. *Useful power* is the rate at which the flow of energy actually accomplishes work. The units for power are units of energy divided by units of time – for example, British thermal units per hour, kilocalories per minute, and joules per second. One joule per second is a *watt* (w). A kilowatt (kw) is 1000 watts, and a megawatt (Mw) is 1,000,000 watts; these are the units we use for power in this book. These units are perfectly applicable to flows of nonelectric as well as electric energy, although you may be accustomed to them only in the context of electricity. Similarly, the kilowatt hour (kwh), denoting the amount of energy that flows in an hour if the rate (power) is 1 kilowatt, makes sense as a unit of energy outside the electrical context. A kilowatt hour is 3600 kilojoules. For example, we can speak of an automobile using energy (in this case, chemical energy stored in gasoline) at a rate of 100 kilowatts (100 kilojoules of chemical energy per second). In an hour of steady driving at this rate of fuel consumption, the automobile uses 100 kJ/sec multiplied by 3600 sec (the number of seconds in an hour) or 360,000 kJ. The same quantity of energy used in a jumbo jetliner would produce a much larger power – say, 180,000 kw – for a much shorter time (2 sec).

A complete set of conversion factors for units of energy and power appears just inside the covers of this book.

[**] A physicist might object at this point that it isn't always useful, or even possible, to discriminate between objects and fields of electric and magnetic force. This level of technicality will not be needed in this book.

[†] The units of force are units of mass multiplied by units of acceleration. One newton is a mass of 1 kg times an acceleration of 1 m per second (sec) per second ($1\ N = 1\ kg\ m/sec^2$). One joule equals 1 newton-meter ($1\ J = 1\ N\text{-}m = 1\ kg\ m^2/sec^2$).

[††] Zero and 100 on the Celsius, or centigrade, scale of temperature correspond to the freezing point and the boiling point of pure water. The conversion between Celsius (C) and Fahrenheit (F) is: degrees F = 1.8 × degrees C + 32.

Within the molten outer core is a *solid inner core*, also composed of iron and nickel, under enormous pressure.

Many of the characteristics of Earth's solid surface are the result of the operation of *tectonic processes*—the motion of great solid segments of the lithosphere, called *plates*, which slide about over the plastically flowing asthenosphere at a rate of a few centimeters per year. Operating over hundreds of millions of years, such motions have apparently produced displacements of thousands of kilometers. The now widely accepted theory of continental drift holds that the present arrangement of the continents arose in this way, beginning with the breakup of the single supercontinent Pangaea about 200 million years ago.[5]

Some of the main tectonic processes, as they continue to work today, are illustrated in Figure 2-2. At *divergent plate boundaries* on the ocean floor, such as the East Pacific and Mid-Atlantic Ridges, adjacent plates move apart and new crust is created in the gap by *magma* (molten rock), which rises from below and then solidifies. This phenomenon is called seafloor spreading. At *convergent plate boundaries*, as along the western edge of South America, one plate may be driven beneath the other into the asthenosphere. Heat generated by the friction in these *subduction zones* melts some of the crustal rock to produce magma, which rises to feed volcanic activity at the surface. Deep-sea trenches, steep mountain ranges, and powerful earthquakes are other characteristics of these zones of violent collision between plates. At a third type of interface between plates, the plates slide past each other, moving parallel to the boundary. These *parallel-plate boundaries* are characterized by earthquakes with large surface displacements; the San Andreas Fault, which produced the great San Francisco earthquake of 1906, marks such a boundary. The principal plate boundaries are indicated on the map in Figure 2-3.

Many other geophysical processes operate simultaneously with the tectonic motions described above to govern the shape and composition of Earth's crust. These processes include mountain-building by uplifting of the crust, the wearing-away of exposed rock surfaces by the actions of wind, rain, ice, and chemical processes (together these effects are called *weathering*), the transport

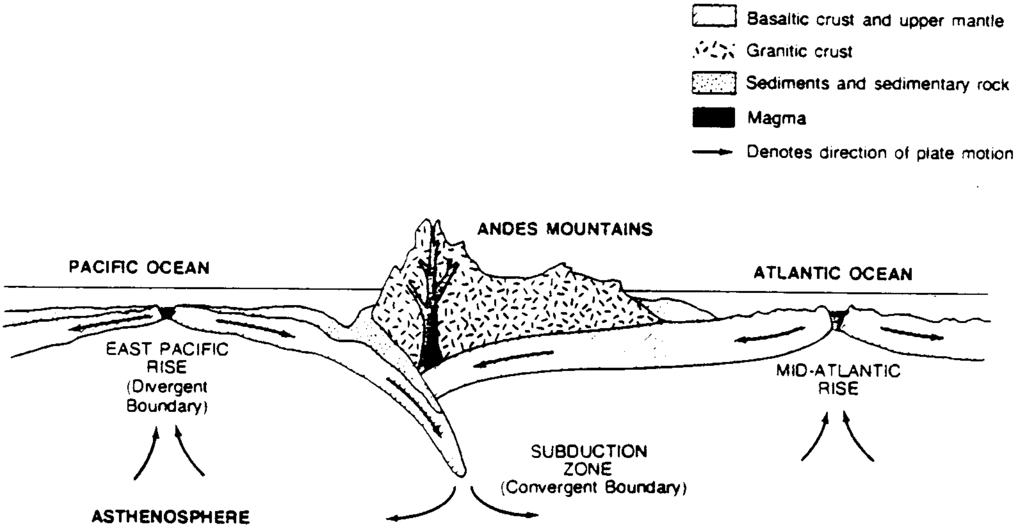


FIGURE 2-2

Tectonic processes and the Earth's surface. (From Rona, 1973.)

---

[5] See J. Tuzo Wilson, ed., *Continents adrift.*

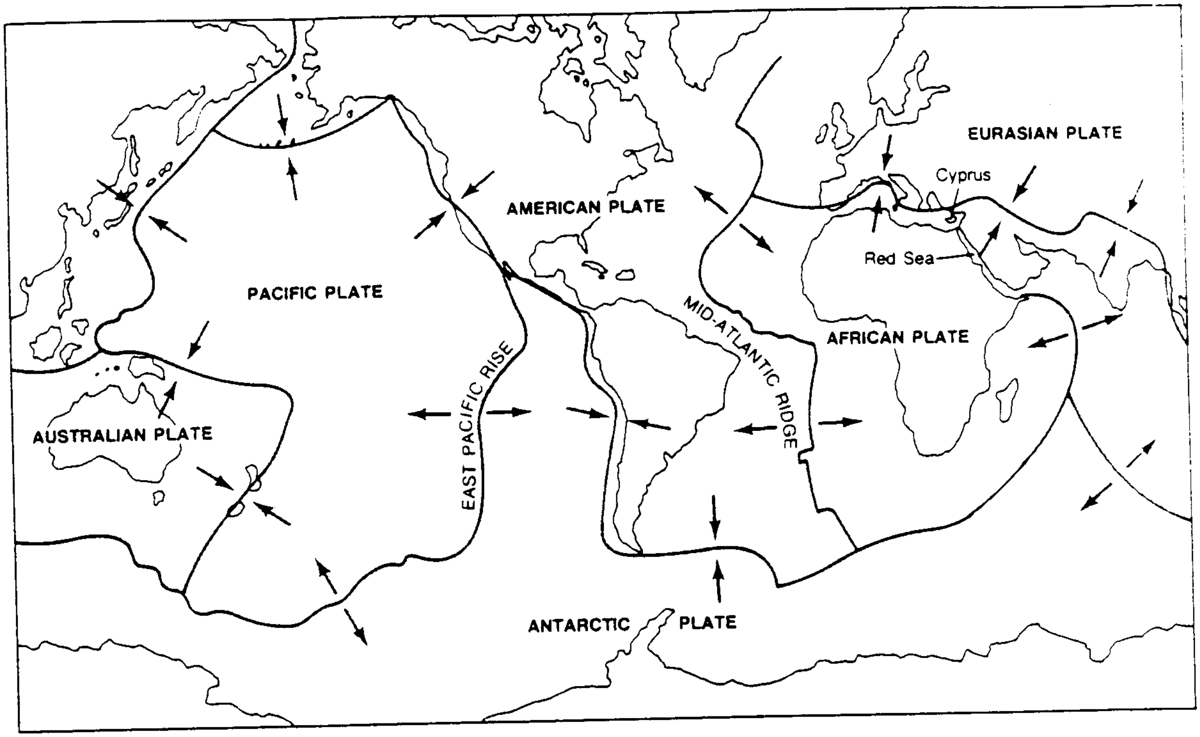


FIGURE 2-3

Six principal tectonic plates of the lithosphere.

of particles of rock and soil by water and wind *(erosion),* and the formation and transformation of new rocks from sedimentary material. The way in which these processes are linked together to produce the principal geological cycles is represented schematically in Figure 2-4.

Rock that is exposed at the surface of the crust is gradually weathered away by physical and chemical processes. The resulting particles are some of the raw materials for new soil (the formation of which also requires the action of living organisms), and some of the chemicals liberated from the rock become available to the biosphere as nutrients (see Chapter 3). Although the rock particles may sometimes be carried uphill by wind and ice, the predominant motion is downhill with the flow of water. Thus it happens over geologic spans of time (hundreds of thousands to millions of years) that large amounts of material are removed from the exposed rocky crust at high elevations and deposited on the lowlands and on the ocean floors. The accumulating weight of these sediments, consisting ultimately not only of rock fragments but also of dead plant and animal matter and chemicals precipitated out of seawater, contributes to sinking of the underlying crust and upper mantle. (Tectonic subsidence, the result of large-scale crustal motions, is more important in this sinking phenomenon than is the local accumulation of sediment weight, however.[6]) Simultaneously, crustal rise, or uplifting, under the lightened regions restores some of the loss of elevation produced by weathering and erosion. This process of sinking and uplifting is made possible by the capacity of the dense but soft (almost molten) asthenosphere to be deformed and, indeed, to flow. The folding

---

[6] Press and Siever, *Earth*, p. 479.

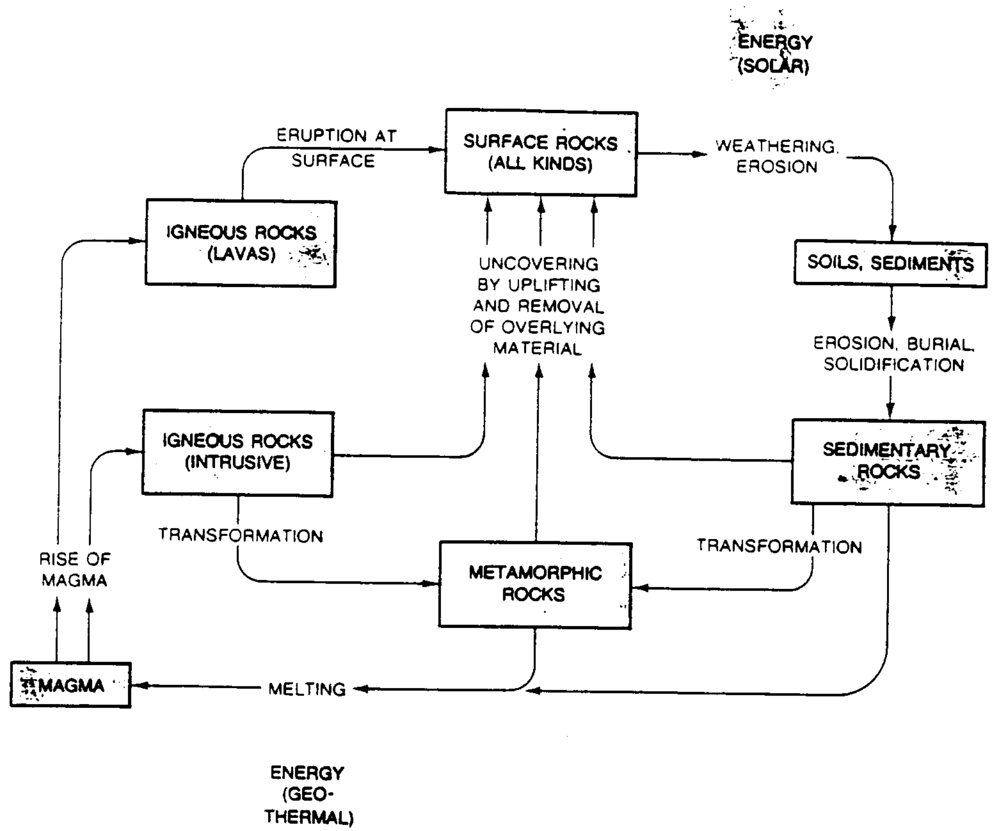


FIGURE 2-4

Geologic cycles. Intrusive igneous rocks are those that solidify from magma before reaching the surface, in contrast to extrusive igneous rocks (lava). Most sedimentation (deposition of sediments) takes place on the ocean floor. The time for material to complete a cycle is typically tens of millions to hundreds of millions of years.

and buckling of Earth's crust, which has produced much the varied topography we see, is the combined result of the sinking-uplifting phenomenon just described and the continuous collision of the great lithospheric plates.

As layers of sediment become more deeply buried, they are subjected to temperatures and pressures high enough to initiate chemical and physical changes that transform the sediments into rock (called *sedimentary rocks*). Among the rocks formed in this way are shale, sandstone, limestone, and dolomite. Under some conditions, such as the particularly energetic geological environment where tectonic plates collide, further transformations under the action of heat and pressure produce *metamorphic rocks*, among which are slate and marble. The most abundant rocks in Earth's crust, however, are *igneous rocks*—those formed by the cooling and solidification of magma. Repeated local melting, migration, and resolidification of the rock in Earth's crust and upper mantle have led over the eons to a general stratification, with the densest material on the bottom and less dense material above. Thus, the upper layer of the continental crust consists largely of *granitic* igneous rocks—rocks rich in the relatively light elements silicon and aluminum. The oceanic crust and the lower layer of the continental crust (Figure 2-1) consist mainly of *basaltic* igneous rocks—somewhat denser material, containing substantial amounts of iron in addition to the lighter elements. The mantle below is *olivine* igneous rock, richer yet in iron and therefore denser than the overlying crust.

The average elemental composition of Earth's crust is given in Table 2-4. The predominance of the light elements is apparent: of the ten most abundant elements—accounting for 99 percent of the mass of the crust—only iron has an atomic number above 25. The crust comprises only about 0.4 percent of the mass of Earth, however. Essentially all the rest resides in the denser and vastly thicker mantle and core (Table 2-2). The composition of the entire planet (Table 2-5), reflects the predominance of iron in those inner layers. Of interest is that carbon, the basic building block of living material, is not among the most abundant elements (it ranks fourteenth in crustal abundance, at 0.032 percent).

TABLE 2-4
Average Composition of Earth's Crust

Element	*Atomic number*	*Percentage by weight*
Oxygen	8	45.2
Silicon	14	27.2
Aluminum	13	8.0
Iron	26	5.8
Calcium	20	5.1
Magnesium	12	2.8
Sodium	11	2.3
Potassium	19	1.7
Titanium	22	0.9
Hydrogen	1	0.14

Source: Brian J. Skinner, *Earth resources.*

TABLE 2-5
Average Composition of Earth (Overall)

Element	*Atomic number*	*Percentage by weight*
Iron	26	34.6
Oxygen	8	29.5
Silicon	14	15.2
Magnesium	12	12.7
Nickel	28	2.4
Sulfur	16	1.9
Calcium	20	1.1
Aluminum	13	1.1
Sodium	11	0.57
Chromium	24	0.26

Source: Brian Mason, *Principles of geochemistry.*

The energy that drives the great geological cycles has two distinct origins. Those parts of the cycles that take place on the surface—weathering, the formation of soil, erosion, the production of plant and animal matter that contributes to sediments—are powered by solar energy and its derivatives, wind and falling water. (The character of these energies is examined more closely later in this chapter.) The remaining geophysical processes (for example, the production and migration of magma and the inexorable motions of the tectonic plates) are driven by geothermal energy—heat that is produced beneath Earth's surface. It is thought that most of this heat results from the decay of radioactive isotopes that were already present when Earth was formed.[7] (The reader completely unfamiliar with the terminology and physics of radioactivity may wish to look ahead to Box 8-3.) The most important isotopes in this respect are uranium-238 (half-life 4.5 billion yr), thorium-232 (half-life 14 billion yr), and potassium-40 (half-life 1.3 billion yr). Notwithstanding the rather low concentration of these isotopes in Earth's crust, the energy released by their continuing radioactive decay is enough to account approximately for the observed rate of heat flow to the surface. The very long half-lives of these isotopes guarantee, moreover, that this source of energy for geological change will have been diminished only slightly a billion years hence.

The processes of melting and resolidification, sinking and uplifting, the motion of tectonic plates of continental scale, the gouging and pushing of massive glaciers, and the different rates of weathering and erosion associated with different climates and different combinations of exposed rocks have combined to produce a tremendous variety of geological features.[8] The importance of these features to human beings is severalfold. The landforms—plains, mountains, valleys, and so on—are one major determinant of the extent to which different parts of the planet's surface are habitable. The soils that have resulted from geological and biological processes over

---

[7] It is possible that there is some additional contribution by frictional heat generation resulting from tidal forces on the molten and plastic parts of Earth's interior.

[8] The reader interested in pursuing this complex but fascinating subject should consult one of the several good geology books listed at the end of this chapter.

the millennia are another (and more limiting) determinant of how many people can be supported and where. The zones of earthquakes and volcanism present serious environmental hazards to humans. And the distribution of fossil fuels and metals in scattered deposits far richer than the average crustal abundance is a geological phenomenon of enormous practical importance.

Although the processes that produced these features often act imperceptibly slowly in human terms, the temptation to consider the geological forces to be beyond human influence – or to take for granted their contributions to human well-being – must be resisted. Soil that has taken a thousand years to accumulate can be washed or blown away in a day through human carelessness; and there is evidence that the activities of human societies worldwide – cultivation, overgrazing, deforestation, construction – have doubled the prehistoric rate of sediment transport to the sea (Chapter 6 and Chapter 11). Earthquakes are widely feared and are called natural disasters, but the lack of foresight in planning and construction that has characterized the development of human settlements in active earthquake zones suggests that the consequences are due as much to human ignorance and irresponsibility as to nature's harshness. There are circumstances, moreover, in which human activities actually *cause* earthquakes (injection of liquid wastes into rock formations, for example), and conceivably there may someday be technological means by which the frequency of strong earthquakes can be diminished. Without concentrated deposits of mineral ores, industrial civilization as we know it could not have arisen; they represent a coincidental natural subsidy for society, provided by the work of natural energy flows over eons. The notion that technological civilization is now clever enough to do without this subsidy, once it is used up, by extracting needed materials from common rock is a dubious one. (We will examine this idea more closely in Chapter 9.)

THE HYDROSPHERE

Most forms of life on Earth require the simultaneous availability of mineral nutrients, certain gases, and water in liquid form. The boundaries of the biosphere – which are fuzzy, rather than sharp – can be defined as the places where the concentration of one or more of these essentials drops too low to sustain life. The principal reservoirs of available mineral nutrients are soil and sediment, the main reservoir of the needed gases is the atmosphere, and the primary supply of water is, of course, in the oceans. Where they meet, these reservoirs intermingle to produce the most fertile parts of the biosphere: the upper layers of soil, where gases and moisture readily penetrate, and the shallower parts of the oceans, where nutrients from the land and the bottom mingle with dissolved gases and light that penetrates downward from the surface.

The oceans include not only some of the planet's most hospitable environments for life (and, almost surely, the environment where life began) but also make up by far the largest single habitat on Earth's surface. They cover almost 71 percent of the planet and their volume is an almost incomprehensible 1.37 billion cubic kilometers (330 million mi^3). The term ***hydrosphere*** refers not only to the oceans themselves, however, but also to the "extensions" of the oceans in other realms – the water vapor and water droplets in the atmosphere, the lakes and the rivers; the water in soil and in pockets deep in layers of rock; the water locked up in ice caps and glaciers. The sections that follow here examine, first, some of the important characteristics of the oceans, then the behavior of ice on Earth's surface, and, finally, the hydrologic cycle, which makes water so widely available even far from the seas.

The Oceans

More than 97 percent of the water on or near the surface of Earth is in the oceans (Table 2-6). This enormous reservoir is a brine (salts dissolved in water) of almost uniform composition. The concentration of the dissolved salts ranges from 3.45 percent (by weight) to about 3.65 percent, varying with depth and latitude. The density of seawater varies between 1.026 and 1.030 grams per cubic centimeter, depending on depth and salinity, compared to 1.000 grams per cubic centimeter for-fresh water at the reference temperature of 4° C (39° F). An average cubic meter[9] of seawater weighs 1027 kilograms

---

[9] One cubic meter (m^3) = 35.3 cubic feet = 264 gallons.

TABLE 2-6
Water Storage in the Hydrosphere

Storage	*Volume (1000 km³)**
Average in stream channels	1
Vapor and clouds in atmosphere	13
Soil water (above water table)	67
Saline lakes and inland seas	104
Freshwater lakes	125**
Groundwater (half less than 800 m below Earth's surface)	8,300
Ice caps and glaciers	29,200
Oceans	1,370,000

*1 km^3 = 264 billion gallons.
**Twenty percent of this total is in Lake Baikal in the Soviet Union.
Source: Brian J. Skinner, *Earth resources.*

TABLE 2-7
*Composition of Seawater (excluding dissolved gases) (MT/km^3)**

Elements at more than 1000 MT/km^3		*Selected elements at less than 1000 MT/km^3*	
H_2O	991,000,000	Lithium	175.0
Chlorine	19,600,000	Phospohorus	70.0
Sodium	10,900,000	Iodine	60.0
Magnesium	1,400,000	Molybdenum	10.0
Sulfur	920,000	Copper	3.0
Calcium	420,000	Uranium	3.0
Potassium	390,000	Nickel	2.0
Bromine	67,000	Cesium	0.4
Carbon	29,000	Silver	0.2
Strontium	8,300	Thorium	0.04
Boron	5,000	Lead	0.02
Silicon	3,100	Mercury	0.02
Fluorine	1,300	Gold	0.004

*1 MT = 1000 kg.
Source: Edward Wenk, Jr., The physical resources of the ocean, p. 167.

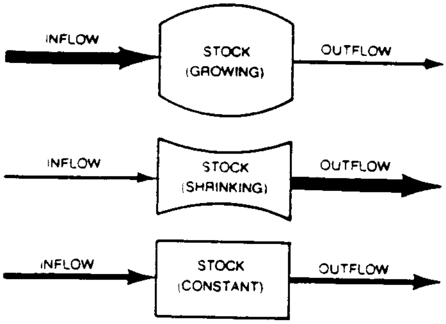
(1.027 MT, or 1.13 short tons), of which about 36 kilograms is dissolved salts. Although most of this material is the familiar sodium chloride, more than half of all the known elements are present in seawater at trace concentrations or more.

The concentrations of several elements in seawater are given in Table 2-7. It is thought that this composition has remained essentially unchanged during most of geologic time. This would mean that the inflow of minerals reaching the oceans from rivers, from the atmosphere, and from undersea volcanoes has been roughly balanced by the outflow—namely, the incorporation of inorganic precipitates and dead organic matter into sediments on the ocean floor.[10] In a situation of *equilibrium* of this kind (with inflow balancing outflow), it is easy to calculate the average time an atom of a given element spends in the ocean between entering it and leaving it. This is called the *residence time;* it is an important concept in the study of nutrient cycles and of pollution. The concepts of equilibrium and residence time, along with some related ideas that find widespread application in environmental sciences, are reviewed in Box 2-2. The residence times of

---

[10] A superb and detailed treatment of the processes maintaining the composition of the oceans appears in Ferren MacIntyre, Why the sea is salt. This and other *Scientific American* articles on the oceans referred to in this section are collected in J. Robert Moore, ed., *Oceanography.*

BOX 2-2 Flows, Stocks, and Equilibrium



The terms *mass balance, energy balance, input/output analysis,* and *balancing the books* all refer to fundamentally the same kind of calculation—one that finds extensive application in physics, chemistry, biology, and economics and in the many disciplines where these sciences are put to use. The basic idea is very simple: everything has to go somewhere, and it is possible and useful to keep track of where and how fast it goes.

The concepts and terminology are illustrated in the diagram here. A stock is a supply of something in a particular place—money in a savings account, water in a lake, a particular element in the ocean. The stocks can be measured in terms of value (dollars), volume (liters), mass (grams), energy (joules), number of molecules, or other units, but not time. Time appears, instead, in the complementary concept of *flows,* the inflow (or input) being the amount of commodity added to the stock per unit of time and the outflow (or output) being the amount of commodity removed from the stock per unit of time. Thus, flows are measured in units like dollars per year, liters per minute, grams per day, or joules per second (watts). In the diagram, the sizes of the flows are indicated by the widths of the arrows. In a savings account, the inflow is deposits plus interest, the outflow is withdrawals, and the stock is the balance at any given time.

Clearly, if the inflow is greater than the outflow, the stock becomes larger as time passes; if outflow exceeds inflow, the stock shrinks. The change in the size of the stock in a given period is the difference between inflow and outflow, multiplied by the length of the period. (If the inflow and outflow vary during the period, one must use their averages.) In the event that the inflow and the outflow have exactly the same magnitude, the size of the stock remains constant. This last situation, where inflow and outflow balance is called equilibrium, or, more specifically, *dynamic equilibrium* (something is flowing, but nothing is changing). The more restrictive case where nothing at all is happening—that is, no inflow, no outflow—is called *static equilibrium.*

In a state of equilibrium, there is not necessarily any relation between the size of the flows (*throughput*) and the size of the stock.* For example, a small lake in equilibrium may be fed by a large river and drained by an equally large one (small stock, large throughput), or a large lake in equilibrium may be fed by a small river and drained by an equally small one (large stock, small throughput). If one divides the size of a stock in equilibrium by the size of the throughput, one obtains a very useful quantity—the *average residence time* (τ). In the example of the lakes, this is the average length of time a water molecule spends in the lake between entering and leaving. For a lake of 100 million cubic meters, fed and drained by two rivers with flows of 100 cubic meters per second each, the average residence time would be given by:

$$\tau = \frac{100{,}000{,}000\ \text{m}^3}{100\ \text{m}^3/\text{sec}} = 1{,}000{,}000\ \text{sec},$$

which is about twelve days. A smaller volume and/or a larger throughput would produce a shorter residence time.

The concept of residence time is useful not only for describing geophysical processes but also for analyzing economic and biological ones. Economic "residence times" include replacement periods for capital and labor. And, to cite an example from biology, if the stock is the world's human population, then the inflow is the rate at which people are born, the outflow is the rate at which people die, and the average residence time is the life expectancy. Clearly, a given population size could be maintained at equilibrium by conditions of high throughput (high birth rate, high death rate, short life expectancy) or by conditions of low throughput (low birth rate, low death rate, long life expectancy). This subject is taken up in more detail in later chapters.

---

*Throughput means just what you would think—"what flows through"—and in general has the magnitude of the smaller of the inflow and outflow in a given situation. In equilibrium, inflow and outflow and throughput are all the same number.

TABLE 2-8
Residence Times of Some Constituents of Seawater

Element	*Residence time (million years)*
Sodium	260
Magnesium	45
Calcium	8
Potassium	11
Silicon	0.01

Source: Strahler and Strahler, *Environmental geoscience*, p. 197.

some important constituents of seawater are listed in Table 2-8.

The absolute quantities of materials dissolved in a cubic kilometer of seawater are quite large – 175 metric tons of lithium, 3 metric tons of copper, 200 kilograms of silver, and 4 kilograms of gold, to mention some elements commonly regarded as scarce. Multiplying such numbers by the total volume of the oceans gives very large numbers, indeed, and these staggering quantities, combined with the ready accessibility of the oceans, have stimulated much discussion of mining seawater for its riches. In mining, it is the *concentration* of the material that counts, however (a subject to which we return in Chapter 9). To get the three tons of copper in a cubic kilometer of seawater, for instance, this desired material must somehow be separated from the billion metric tons of other elements mixed up with it, and this is not easy.

Probably a more important reason for looking into the details of the composition of the oceans is to evaluate the seriousness of various kinds of ocean pollution. If pollutants such as lead and mercury, for example, are added to the oceans in sufficient quantities to alter substantially the natural concentrations of those elements over large areas, one might suspect that significant biochemical consequences could result. If, on the other hand, the discharge of an element into the ocean produces concentration changes small compared to natural variations in space and time of the concentration of this substance, little or no harm would be expected.

One cannot assume, of course, that substances added to the oceans are quickly diluted by this vast volume of water. How far? How deep? How fast? are the questions, and the answers are found in the rather complicated patterns of horizontal circulation and vertical mixing in the oceans, as well as in the functioning of biological systems that may concentrate them (Chapter 4). Vertical mixing is rapid only near the surface. The turbulence (violent mixing motions) produced by wave action at the surface penetrates only to a depth of from 100 to 200 meters, and this defines the thickness of the layer within which most of the absorbed solar energy is distributed.

Below this warm, well-stirred surface layer is a transitional region called the *thermocline*, where the temperature drops rapidly. In this region there is usually less vertical motion than in the surface layer. Here heat penetrates partly by conduction (molecules passing on energy by jostling their neighbors) but mostly by convection (transport of energy by bulk motion of a warm medium) in large, slow eddies. There are two reasons for the relative lack of vertical motion in the thermocline: (1) motions originating at the surface have been damped out by friction before penetrating so deep and (2) the colder water near the bottom of the thermocline tends to be denser than the warmer water near the top of this layer, stifling thermal circulation.[11] In some circumstances, however, variations in salinity can influence density enough to produce a vertical circulation, despite the countervailing influence of the temperature profile. The bottom of the thermocline lies between 1000 and 1500 meters below the surface; from this level down, the temperature is nearly uniform and lies in the range from 0° to 5° C. (Seawater freezes at −2° C, or about 28.4° F.) Like the thermocline, this deepest ocean layer is thermally stratified, with the coldest water lying at the bottom. In the deep layer, moreover, salinity increases with depth; the saltier water is, the denser it is, so the salinity profile and the temperature profile both place the densest water at the bottom, inhibiting vertical mixing.

This simplified view of the vertical layering of the oceans is illustrated schematically in Figure 2-5. Note that the stratification breaks down near the poles, where

---

[11] Thermal circulation, or thermal convection, occurs when warm fluid (liquid or gas), finding itself below colder, denser fluids, rises, while the colder material sinks. This is what happens when a fluid, such as water in a pan, is heated from below. The ocean is heated mainly at the top

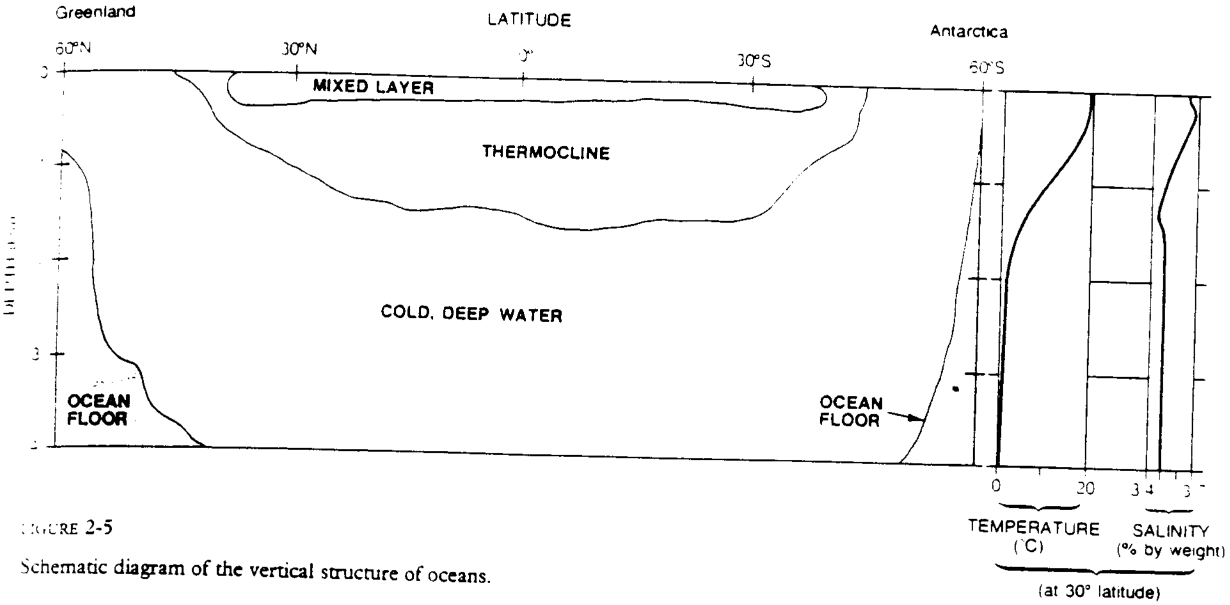


FIGURE 2-5

Schematic diagram of the vertical structure of oceans.

the cold layer extends all the way to the surface. That the surface waters in the Arctic and Antarctic oceans should be considered outcroppings of the deep layer that extends throughout the world ocean is suggested not only by patterns of temperature and salinity but also by the distribution of certain creatures. The huge Greenland shark, for example, once thought to inhabit only Arctic waters, has been photographed three or four kilometers deep in waters off Baja California and in the Indian Ocean.[12]

The stable stratification of the oceans also breaks down at scattered places and times far from the poles, as in upwellings in which winds push the surface water away from a steep continental slope and cold water rises from below to replace it (off the coast of Peru, for example) or when rapid cooling of surface water under unusual circumstances causes it to sink. The mean residence times for water in the various ocean layers illustrate the relative rarity in space and time of large vertical movements: a typical water molecule in the mixed layer may spend 10 years there, whereas one in the thermocline spends 500 years before reaching the deep layer, and a water molecule in the deep layer typically spends 2000 or 3000 years before reaching one of the upper layers. Clearly, it must be assumed that most substances added to the oceans near the surface and dissolved there will remain near the surface for years, being diluted only by the small fraction of the ocean water that makes up the well mixed top layer (between 3 and 5 percent).

Horizontal circulation in the oceans is considerably faster than vertical mixing. Water in the main currents, which generally involve only the mixed layers, typically moves at speeds of 1 kilometer per hour (km/hr), and occasionally up to 5 kilometers per hour. Thus, an object or a substance being carried in the current might easily move 1000 kilometers in a month and cross an ocean in six months to a year. (The main oceanic surface currents appear on the map in Figure 2-6.) The principal features are the circular movements, called *gyres,* centered in the subtropical latitudes (25° to 30° north and south of the equator), an equatorial countercurrent (most prominent in the Pacific) that provides a return pathway for water that would otherwise be piling up against Asia and Australia, and the Antarctic circumpolar current, flowing uninterrupted from west to east around the far southern part of the globe. These currents are produced

---

[12] John D. Isaacs, The nature of oceanic life.

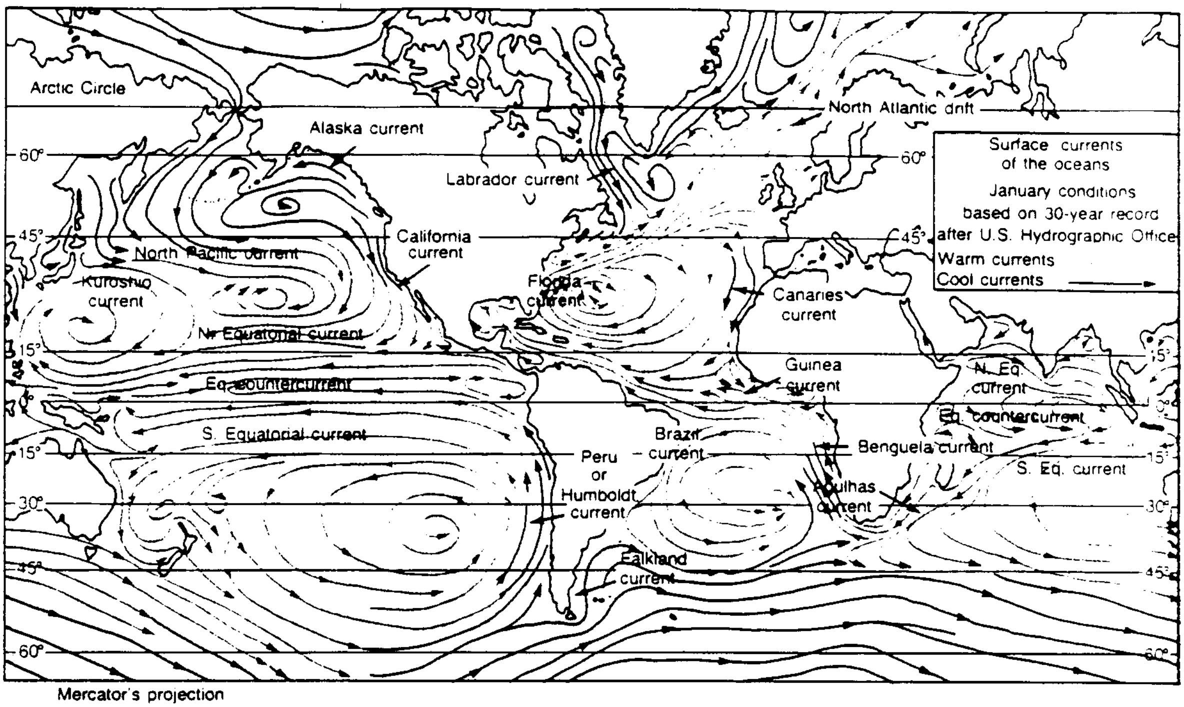


FIGURE 2-6

Main oceanic surface currents.

by a complex interaction of the effects of the winds, Earth's rotation, and the placement of continents and islands.[13] By moving enormous quantities of water—sometimes warm, sometimes cold—from one region to another, the ocean currents exert a major influence on climate, a subject taken up in more detail later. The two mightiest currents on Earth, the Antarctic circumpolar current and the main branch of the Gulf Stream, each carries some fifty times the combined flow of all the world's rivers.[14]

The horizontal circulation in the deep layer of the ocean is much less thoroughly mapped than that of the surface layer and has been widely supposed to be much less vigorous. Typical speeds in these deep currents have been thought to be on the order of 0.1 kilometer per hour or less. An increasing number of direct measurements of deep ocean currents now suggest a much more vigorous deep-ocean circulation, however, involving powerful eddies 100 kilometers or more in horizontal extent, containing currents of 0.5 to 1 kilometer per hour.[15] One would expect the general flow to be from the poles toward the equator, inasmuch as some cold water enters the deep layer from above at the poles and some rises in upwellings closer to the equator. But the actual situation is made quite complicated by Earth's rotation, by the irregular distribution of landmasses, and by the complex topography of the ocean floor.

It has been true historically and is still true today that the usefulness of the oceans to civilization and, in turn, civilization's impact on the oceans have been greatest in the shallower waters at the edges of the continents. This

---

[13] See, for example, R. W. Stewart, The atmosphere and the ocean.

[14] P. H. Kuenen, *Realms of water*, p. 47.

[15] F. Bretherton, Recent developments in dynamical oceanography.

is so partly for simple reasons of accessibility, and partly because of the particular fertility of the near-shore waters and the richness of the underlying sediment in minerals of economic interest. The term *continental shelf* refers to that part of the near-shore underwater topography that is actually an extension of flatlands on a continent itself. Although the outer edge of a continental shelf is often defined as the line along which the depth of the water reaches 200 meters, a less arbitrary boundary is the point where there is a marked increase in the downward slope of the bottom. The steeply sloping region just beyond this boundary is called the *continental slope.* The "foothills" leading from the ocean floor to the seaward edge of the continental slope are called the *continental rise.*

Using the definition of continental shelf just given, it has been estimated that continental shelves underlie 7.5 percent of the area of the oceans (an area equal, however, to 18 percent of Earth's land area).[16] These shelves vary in width from essentially nothing to 1500 kilometers, and their seaward edges vary in depth from 20 to 550 meters (the average depth at the edge is 133 meters). The circulation patterns in shallow, continental-shelf waters are complex, and the residence times of dissolved substances over the shelf can be surprisingly long—as much as several years to migrate from the coastline to the outer edge of a wide shelf like that off the east coast of the United States.

Glaciers and Sea Ice

Fifteen thousand years ago, much of what is now continental shelf was dry land. Sea level was 130 meters lower than it is today. Where was the 45 million cubic kilometers of water (about 12 billion billion gallons!) this difference in sea level represents? It was locked up in the great glaciers of the ice age. In the warmer period in which we find ourselves today, the water that remains frozen as ice still far exceeds all other reservoirs of fresh water on Earth (Table 2-6). Were this ice to melt, sea level would rise another 80 meters.

It is important in this connection to distinguish between *glaciers* and *sea ice.* A glacier is a sheet of ice formed *on land* when accumulated snow is compressed and hardened into ice by the weight of overlying layers. Sea ice is ice formed from seawater; it floats on the ocean's surface, although it may be attached to land at its edges.

The glaciers that usually come to mind when one hears this term are the scattered "mountain and valley glaciers" that occur throughout the world's high mountain ranges—the Himalayas, Andes, Rockies, and Alps, for example. The larger glaciers of this variety are some tens of kilometers long, a kilometer or more across, and a few hundred meters thick. These glaciers are constantly in motion, being fed by snowfall at their surfaces in the higher elevations and moving downhill as the deep layers, under great pressure, flow as a plastic solid. (Perhaps the easiest way to visualize what is going on in such flow, which as we noted earlier also occurs in rock in Earth's mantle, is to consider the plastic solid to be an extremely viscous fluid.) The speed of advance varies along the length of the glacier, but is typically 100 meters per year or more in the main body of the larger mountain glaciers. The advance of such glaciers is terminated by melting of the tongue of the glacier at the lower end.

By far the greatest part of the world's inventory of ice—more than 99 percent—is tied up in a second kind of glacier, the land ice or ice sheets, that cover the bulk of the Greenland and Antarctic landmasses. The formation of such a sheet requires an arctic climate, sufficient precipitation, and fairly flat land. The ice layer that results covers essentially the entire landscape in a gently sloping dome, interrupted only by a few projecting mountain peaks. The Greenland ice sheet covers an area of 1.74 million square kilometers (about 80 percent of the total area of Greenland) and has an average thickness of about 1600 meters (5250 ft). The Antarctic ice sheet covers 13 million square kilometers with ice up to 4000 meters thick (13,000 ft) and averaging perhaps 2300 meters.[17] About 91 percent of the world's ice thus is in the Antarctic sheet and about 9 percent in the Greenland sheet. These ice sheets, like mountain and valley glaciers,

---

K. O Emery, The continental shelves.

[17] Press and Siever, *Earth,* p. 371; Strahler and Strahler, *Environmental geoscience,* pp. 434–436.

are in motion, carrying to the sea the ice formed from precipitation in the central regions. Typical speeds are some tens of meters per year on the ice sheet proper, but they can be much higher—hundreds, and even thousands, of meters per year—where certain glacial tongues meet the sea.

Where the ice sheets meet the sea in broad expanses, they may extend into the ocean as more-or-less floating ice shelves, from tens to hundreds of meters thick. In the Antarctic, these shelves reach widths of hundreds of kilometers. The largest, the Ross Ice Shelf, covers more than 500,000 square kilometers. Icebergs originate when great masses of ice break off from the tips of glacial tongues or the edges of ice shelves and are carried away (often into shipping lanes) by currents (see Figure 2-7).

Sea ice, as distinguished from floating extensions or pieces of glaciers, is formed by the freezing of seawater on the ocean surface. The North Pole ice pack, with a mean extent of about 10 million square kilometers, is a collection of slabs of sea ice floating on the Arctic Ocean. In winter, these slabs are frozen together and attached to land at various points around the ocean's periphery. In summer, some of the slabs break apart and are separated by narrow strips of open water, and the southern limit of the ice retreats northward. The sea ice, which begins to form at − 2° C, is porous, and the enclosed cavities often contain water saltier than seawater. Glacial ice, by contrast, consists of fresh water, being simply compacted and recrystallized snow.

The maximum thickness of sea ice is only between 3 and 5 meters. Once it reaches this thickness, the layer of ice insulates the underlying water so well that no more can freeze—heat is supplied from the deeper water faster than the surface layer can lose it through the ice. (Ice is a poor conductor of heat and snow an even poorer one, which is why snow igloos stay so warm inside.) If the average thickness of the North Pole ice pack is 2 meters, it contains less than one-hundredth as much ice as the Greenland ice sheet.[18] Of course, melting of the sea ice would have no direct effect on sea level, even if the volume of this ice were much greater; the ice is floating, thus displacing an amount of water equal to its weight, so

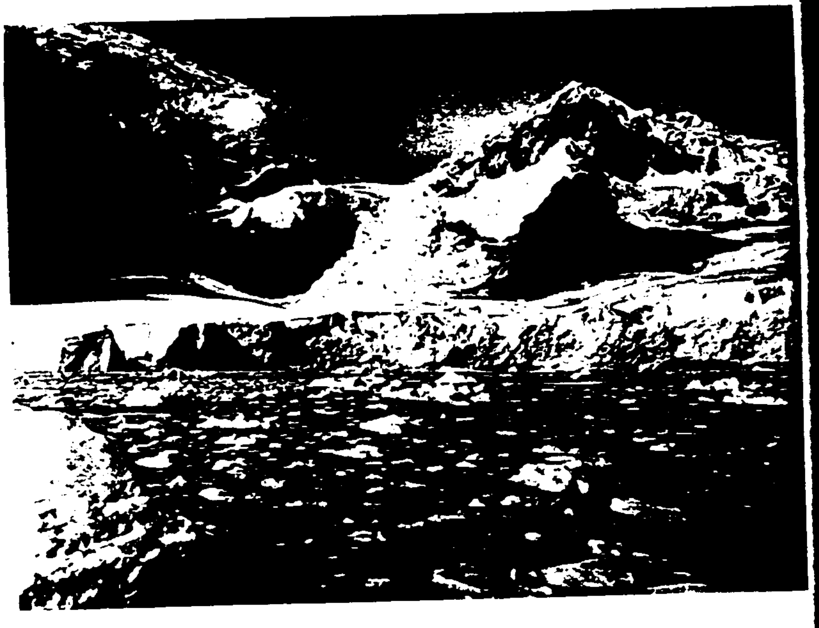


FIGURE 2-7

A glacier feeding ice into the sea in Paradise Bay, Antarctica. No large icebergs are visible in this picture. (Photo by P. R. Ehrlich.)

---

[18] Ten million km^2 of area multiplied by 0.002 km average thickness is 20,000 km^3 of ice in the polar pack, compared to 1.7 million km^2 multiplied by 2.2 km average thickness, or 2,700,000 km^3 of ice, in the Greenland sheet.

it is already contributing exactly as much to the level of the oceans as it would if it melted. Aside from the sea-level issue, however, which relates solely to glacial ice sheets, the sea ice has great importance for climate.

The Hydrologic Cycle

Although oceans and ice caps contain some 99.3 percent of all the water on Earth (Table 2-6), the fraction of 1 percent residing at any given time in the atmosphere, in lakes and streams, and in soil and subsurface layers plays unique and important roles. The flow of water on the surface is a major determinant of the configuration of the physical environment. Soil moisture is essential to most terrestrial plant life. The stocks and flows of ground and surface water are major links in the transport and cycling of chemical nutrients and important determinants of what kinds and intensities of human activity can be supported in what locations. And water in the atmosphere has several functions that are central to shaping climates.

The set of processes that maintain the flow of water through the terrestrial and atmospheric branches of the hydrosphere is called the *hydrologic cycle*. The cycle includes all three physical states of water—liquid, solid (ice and snow), and gas (water vapor). It also includes all of the possible transformations among these states—*vaporization*, or evaporation (liquid to gas); *condensation* (gas to liquid), *freezing* (liquid to solid); melting, or *fusion* (solid to liquid); and *sublimation* (gas to solid, or the reverse).

The principal flows in the hydrologic cycle are: (1) *evaporation* of water from the surface of the oceans and other bodies of water, and from the soil; (2) *transpiration* of water by plants, the result of which is the same as that of evaporation—namely, the addition of water vapor to the atmosphere; (3) horizontal *transport* of atmospheric water from one place to another, either as vapor or as the liquid water droplets and ice crystals in clouds; (4) *precipitation*, in which atmospheric water vapor condenses (and perhaps freezes) or sublimates and falls on the oceans and the continents as rain, sleet, hail, or snow; and (5) *runoff*, in which water that has fallen on the continents as precipitation finds its way, flowing on and under the surface, back to the oceans. Because it is difficult and not particularly useful to distinguish between the contributions of evaporation and transpiration on the continents, these two terms are often lumped together as *evapotranspiration*.

The magnitudes of these flows, averaged over all the continents and oceans and expressed in thousands of cubic kilometers of water per year, are shown in Figure 2-8.[19] These magnitudes are based on the assumption that the various components of the hydrosphere are in equilibrium, which is at least a good first approximation. That is, on a year-round average, inflows and outflows for the atmosphere, the oceans, and the continents all balance. (For example, in thousands of cubic kilometers, the atmosphere receives 62 + 456 = 518 as evaporation from the surface and gives up 108 + 410 = 518 as precipitation.)

The magnitude of the flows in the hydrologic cycle is more readily grasped if one thinks of the flows in terms of the equivalent depth of water, averaged over the surface area involved. In these terms, the world's oceans annually lose to evaporation a layer of water 1.26 meters deep (about 4 feet) over their entire surfaces, gaining back 1.14 meters from precipitation and 0.12 meters from the discharge of rivers and groundwater. The continents receive precipitation each year equivalent to a layer of water 0.73 meters (29 in) deep over their entire surface areas, of which 0.42 meters is lost to evaporation and 0.31 meters makes up the runoff.

Combining the foregoing information on equilibrium flows with the information on stocks in the hydrosphere summarized in Table 2-6 permits us to estimate the average residence time of water in the different parts of the cycle (see Box 2-2). These residence times, which are of great importance in analyzing the transport of pollutants, as well as nutrients, by the hydrologic cycle, are listed in Table 2-9. There is an enormous range, from the average nine days a water molecule spends in the atmosphere between being evaporated from the surface and falling again as precipitation, to the 10,000 years a

---

[19] The values presented here are at the high end of a range of published estimates, in which the differences of professional opinion amount to as much as 25 percent.

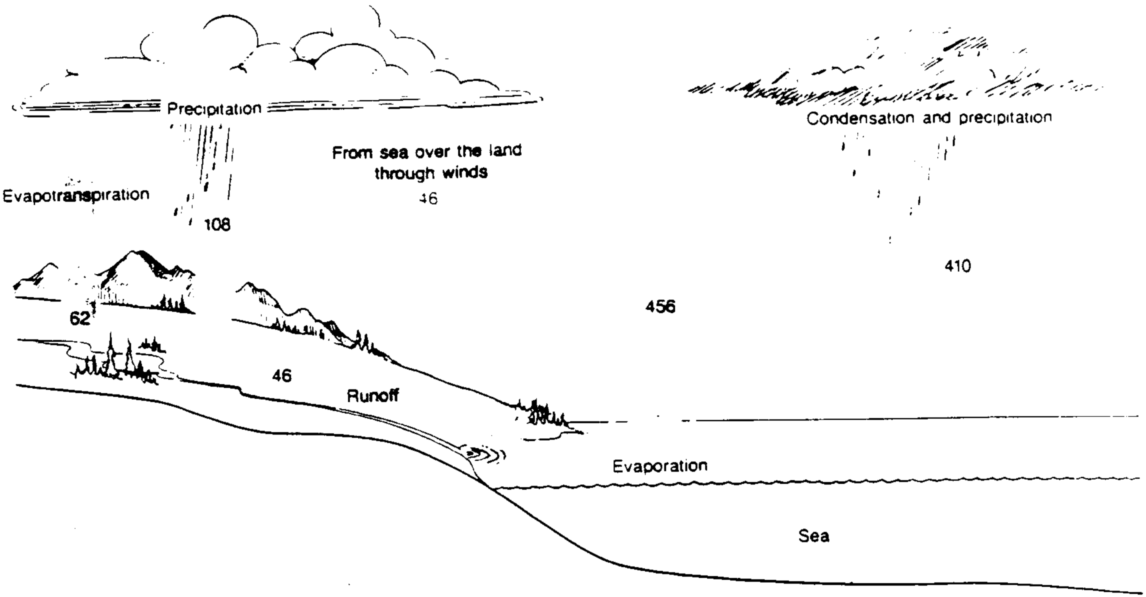


FIGURE 2-8

The hydrologic cycle (1000 km³/yr). (Data from M. I. Budyko, 1974.)

TABLE 2-9
Residence Times of Water Molecules in the Hydrologic Cycle

Location	*Residence time*
Atmosphere	9 days
Rivers (typical speed, 1 m/sec)	2 weeks
Soil moisture	2 weeks to 1 year
Largest lakes	10 years
Shallow groundwater (speed, 1–10 m/day)	10s to 100s of years
Mixed layer of oceans (1st 150 m)	120 years
World ocean	3000 years
Deep groundwater	up to 10,000 years
Antarctic ice cap	10,000 years

Sources: Computed from Table 2-6 and Figure 2-8 or adapted from SMIC, *Inadvertent climate modification*, and Strahler and Strahler, *Environmental geoscience*.

molecule of water typically spends as ice between falling in a snowflake on the Antarctic ice sheet and rejoining the ocean with the melting of an iceberg. It is also important to remember that there are large deviations from the average in any given category – a water molecule may fall in a raindrop not nine days but an hour after being evaporated from Earth's surface; another may wander not two weeks but two years in the delta of the Amazon River before reaching the sea. Nevertheless, the average residence times can provide useful insights into a variety of important problems, and the approach can be refined whenever information more pertinent than global averages is available.

The balance between precipitation and evapotranspiration varies widely from continent to continent, as shown in Table 2-10. The size of the runoff (the difference between precipitation and evapotranspiration) is a measure of how much water is potentially available for domestic and industrial uses by society (including dilution and removal of wastes) and for the other functions that flowing water performs. Note in Table

TABLE 2-10
Average Water Balance of the Continents

	Precipitation (cm/yr)	Evaporation (cm/yr)	Runoff	
			(cm/yr)	(km³/yr)
Africa	69	43	26	7,700
Asia*	60	31	29	13,000
Australia	47	42	5	380
Europe	64	39	25	2,200
North America	66	32	34	8,100
South America	163	70	93	16,600

*Includes entire USSR.
Source: Budyko, M. I., p. 227.

2-10 the remarkable fact that South America has a runoff per unit of surface area almost three times that of North America, the continent with the next greatest runoff. It is perhaps not so surprising, then, that the discharge of the Amazon River, which drains the wettest third of South America, amounts to about a seventh of the runoff of the entire world.

Much of the runoff on the continents takes place not on the surface but beneath it. Although the quantities can only be estimated, it is clear that most rivers receive at least as much of their flow from seepage through the ground as from flow over the ground; and a certain amount of water reaches the oceans via flowing aquifers and seepage at the edges of the continents without ever joining a surface river at all.[20] Water beneath the land's surface is called *soil moisture,* or soil water, when it is distributed in the first meter or so of soil (a zone defined by the depth of penetration of the roots of most plants.) Below the zone of soil moisture is an intermediate zone where the water percolates downward through open pores in the soil and rock; and below this is the *water table,* marking the surface of the body of *groundwater* that saturates the soil or rock in which it finds itself, filling all pores and spaces in the soil or rock completely. The groundwater extends downward until it is limited by an impermeable layer of rock. In some circumstances, there are successive layers of groundwater (*aquifers*) separated by impermeable layers of rock. The absolute lowest limit of groundwater is probably about 16 kilometers from the surface, where the pressure is so great that all pores are closed and any rock becomes impermeable.

Most groundwater is flowing, albeit very slowly [10 meters per day (m/d) in coarse gravel near the surface, more commonly 1 m/d, and much more slowly at greater depths.] At 1 meter per day, of course, it takes almost three years to move 1 kilometer. Figure 2-9 is a schematic diagram of the zones and flows of subsurface water, showing the intersection of the water table and a surface river.

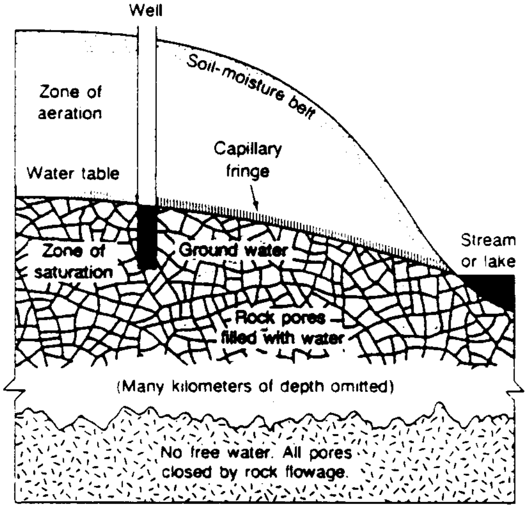


FIGURE 2-9

Zones of subsurface water. (After Ackerman, Colman, and Ogrosky, in A. N. Strahler and A. H. Strahler, *Environmental geoscience.*)

See, for example, Strahler and Strahler, *Environmental geoscience,* Chapter 12, and Kuenen, *Realms of Water,* Chapter 5.

The energy that drives the hydrologic cycle is energy from the sun – indeed, this function is the largest single user of the solar energy reaching Earth's surface. The reason so much energy is required is that it takes a great deal of energy to evaporate water – 2250 joules per gram at the boiling point of 100° C and 2440 joules per gram at Earth's average surface temperature of 15° C. (This is the highest heat of vaporization of any known substance.) It takes fifty times as much energy to evaporate a gram of water as it does to lift it to an altitude of 5 kilometers. The energy used to evaporate the water is stored as latent heat of vaporization (see Box 2-1), which is released to the environment as heat whenever and wherever the water vapor condenses into liquid. Thus, energy delivered by the sun at one point on Earth's surface may be released high in the atmosphere over a point 1000 kilometers away. This mechanism of redistributing energy by the transport and condensation of water vapor is a major determinant of Earth's climate.

As noted above, the energy the sun supplies at the time of evaporation reappears as heat at the time of condensation. Similarly, the smaller amount of solar energy that does the work of lifting the water vapor against the force of gravity appears as frictional heat when falling droplets of condensed vapor collide with molecules of air and when rushing mountain streams rub against their rocky beds. That all the energy the sun supplies to terrestrial processes comes back again in one form or another is not coincidence or quirk, but an illustration of the first law of thermodynamics – the law of conservation of energy. Further excursions into the machinery of the physical world – and of human technology – will require some familiarity with this law and with its companion, the second law of thermodynamics, so an introduction to both is provided in Box 2-3.

ATMOSPHERE AND CLIMATE

The blanket of gases that makes up Earth's atmosphere has many functions. Of the four elements required in greatest quantity by living organisms [carbon (C), oxygen (O), hydrogen (H), nitrogen (N)] the atmosphere provides the main reservoir of one (N), the most accessible reservoir of two others (C, O), and an essential link in the continuous recycling of the fourth (H, in the form of H_2O). The atmosphere is substantial enough to protect the organisms on Earth's surface from a variety of harmful particles and radiations that reach the planet from the sun and from space, but it is transparent enough to permit an adequate amount of life-giving sunlight to penetrate to that surface. Acting as a thermal insulator, the atmosphere keeps Earth's surface much warmer, on the average, than it would be if there were no atmosphere. And the stirrings of the atmosphere, transporting energy and moisture from one place to another, are a major part of the patterns of climate so important to the character and distribution of life.

For simplicity, we begin our investigation of the atmosphere by ignoring its internal vertical and horizontal motions and considering its properties as a static body of gas.

Air

The term *air* refers to the particular mixture of gaseous compounds making up the atmosphere. The average composition of this mixture, not including water, is shown in Table 2-11. An important property of gases is that a given number of molecules at a given temperature and pressure will occupy almost exactly the same volume, regardless of the mass or size of the molecules. This property has led to the use of the somewhat confusing terms *percent by volume* or *fraction by volume* to describe the relative abundance of the various constituents of gaseous mixtures. That is, if three-quarters of the molecules in a container of fixed volume are gas *A* and one-quarter are gas *B*, one can think of gas *A* as "occupying" three-fourths of the volume and gas *B* as "occupying" one-fourth. (What is happening in reality, of course, is that both gases, mixed together, occupy the whole volume, with gas *A* accounting for three-fourths of the pressure in the volume and gas *B* accounting for one-fourth of the pressure.) We will use the term *molecular fraction* (number of molecules of a constituent divided by the total number of molecules in the mixture), because it is unambiguous and works for solids and liquids as well as gases. The reader should simply be aware that this term is interchangeable with the term

TABLE 2-11

Average Composition of Clean Dry Air

Constituent	*Symbol*	*Molecular weight*	*Molecular fraction of air*	*Mass fraction of air*
Nitrogen	N_2	28	0.7809	0.755
Oxygen	O_2	32	0.2095	0.232
Argon	Ar	40	0.0093	0.013
Carbon Dioxide	CO_2	44	320 ppm	486 ppm
Neon	Ne	20	18 ppm	12 ppm
Helium	He	4	5.2 ppm	0.7 ppm
Methane	CH_4	16	2.9 ppm	1.6 ppm
Krypton	Kr	84	1.1 ppm	3.2 ppm
Nitrous Oxide	N_2O	44	0.5 ppm	0.8 ppm
Hydrogen	H_2	2	0.5 ppm	0.03 ppm
Ozone	O_3	48	0.01 ppm	0.02 ppm

Source: Garrels, Mackenzie, and Hunt, *Chemical cycles.*

BOX 2-3 Availability, Entropy, and the Laws of Thermodynamics

Many processes in nature and in technology involve the transformation of energy from one form into others. For example, light from the sun is transformed, upon striking a meadow, into thermal energy in the warmed soil, rocks, and plants; into latent heat of vaporization as water evaporates from the soil and through the surface of the plants; and into chemical energy captured in the plants by photosynthesis. Some of the thermal energy, in turn, is transformed into infrared electromagnetic radiation heading skyward. The imposing science of thermodynamics is just the set of principles governing the bookkeeping by which one keeps track of energy as it moves through such transformations. A grasp of these principles of bookkeeping is essential to an understanding of many problems in environmental sciences and energy technology.

The essence of the accounting is embodied in two concepts known as the first and second laws of thermodynamics. No exception to either one has ever been observed. The first law, also known as the law of conservation of energy, says that energy can neither be created nor destroyed. If energy in one form or one place disappears, the same amount must show up in another form or another place. In other words, although transformations can alter the *distribution* of amounts of energy among its different forms, the *total* amount of energy, when all forms are taken into account, remains the same. The term *energy consumption,* therefore, is a misnomer; energy is used, but it is not really consumed. One can speak of fuel consumption, because fuel, as such, does get used up. But when we burn gasoline, the amounts of energy that appear as mechanical energy, thermal energy, electromagnetic radiation, and other forms are exactly equal all together to the amount of chemical potential energy that disappears. The accounts must always balance; apparent exceptions have invariably turned out to stem from measurement errors or from overlooking categories. The immediate relevance of the first law for human affairs is often stated succinctly as, "You can't get something for nothing."

Yet, if energy is stored work, it might seem that the first law is also saying, "You can't lose!" (by saying that the total amount of stored work in all forms never changes). If the amount of stored work never diminishes, how can we become worse off? One obvious answer is that we can become worse off if energy flows to places where we can no longer get at it—for example, infrared radiation escaping from Earth into space. Then the stored work is no longer accessible to us, although it still exists. A far more fundamental point, however, is that *different kinds of stored work are not equally convertible into useful, applied work.* We can therefore become worse off if energy is transformed from a more convertible form to a less convertible one, even though no energy is destroyed and even if the energy has not moved to an inaccessible place. The degree of convertibility of energy—stored work—into applied work is often called *availability.*

Energy in forms having high availability (that is, in which a relatively large fraction of the

(*Continued*)

stored work can be converted into applied work) is often called high-grade energy. Correspondingly, energy of which only a small fraction can be converted to applied work is called low-grade energy, and energy that moves from the former category to the latter is said to have been degraded. Electricity and the chemical energy stored in gasoline are examples of high-grade energy; the infrared radiation from a light bulb and the thermal energy in an automobile exhaust are corresponding examples of lower-grade energy. The quantitative measure of the availability of thermal energy is temperature. More specifically, the larger the *temperature difference* between a substance and its environment, the more convertible into applied work is the thermal energy the substance contains; in other words, the greater the temperature difference, the greater the availability. A small pan of water boiling at 100° C in surroundings that are at 20° C represents considerable available energy because of the temperature difference; the water in a swimming pool at the same 20° C temperature as the surroundings contains far more total thermal energy than the water in the pan, but the availability of the thermal energy in the swimming pool is zero, because there is no temperature difference between it and its surroundings.

With this background, one can state succinctly the subtle and overwhelmingly important message of the second law of thermodynamics: *all physical processes, natural and technological, proceed in such a way that the availability of the energy involved decreases.* (Idealized processes can be constructed theoretically in which the availability of the energy involved stays constant, rather than decreasing, but in all real processes there is *some* decrease. The second law says that an *increase* is not possible, even in an ideal process.) As with the first law, apparent violations of the second law often stem from leaving something out of the accounting. In many processes, for example, the availability of energy in some *part* of the affected system increases, but the decrease of availability elsewhere in the system is always large enough to result in a net decrease in availability of energy overall. What is consumed when we use energy, then, is not energy itself but its availability for doing useful work.

The statement of the second law given above is deceptively simple; whole books have been written about equivalent formulations of the law and about its implications. Among the most important of these formulations and implications are the following:

1. In any transformation of energy, some of the energy is degraded.
2. No process is possible whose sole result is the conversion of a given quantity of heat (thermal energy) into an equal amount of useful work.
3. No process is possible whose sole result is the flow of heat from a colder body to a hotter one.
4. The availability of a given quantity of energy can only be used once; that is, the property of convertibility into useful work cannot be "recycled."
5. In spontaneous processes, concentrations (of anything) tend to disperse, structure tends to disappear, order becomes disorder.

That Statements 1 through 4 are equivalent to or follow from our original formulation is readily verified. To see that statement 5 is related to the other statements, however, requires establishing a formal connection between order and availability of energy. This connection has been established in thermodynamics through the concept of *entropy,* a well defined measure of disorder that can be shown to be a measure of unavailability of energy, as well. A statement of the second law that contains or is equivalent to all the others is: *all physical processes proceed in such a way that the entropy of the universe increases.* (Not only can't we win – we can't break even, and we can't get out of the game!)

Consider some everyday examples of various aspects of the second law. If a partitioned container is filled with hot water on one side and cold water on the other and is left to itself, the hot water cools and the cold water warms – heat flows from hotter to colder. Note that the opposite process (the hot water getting hotter and the cold getting colder) does not violate the first law, conservation of energy. That it does not occur illustrates the second law. Indeed, many processes can be imagined that satisfy the first law but violate the second and therefore are not expected to occur. As another example, consider adding a drop of dye to a glass of water. Intuition and the second law dictate that the dye will spread, eventually coloring all the water – concentrations disperse, order (the dye/no dye ar-

BOX 2-3 (*Continued*)

rangement) disappears. The opposite process, the spontaneous concentration of dispersed dye, is consistent with conservation of energy but not with the second law.

A more complicated situation is that of the refrigerator, a device that certainly causes heat to flow from cold objects (the contents of the refrigerator—say, beer—which are made colder) to a hot one (the room, which the refrigerator makes warmer). But this heat flow is not the *sole* result of the operation of the refrigerator: energy must be supplied to the refrigeration cycle from an external source, and this energy is converted to heat and discharged to the room, along with the heat removed from the interior of the refrigerator. Overall, availability of energy has decreased, and entropy has increased.

One illustration of the power of the laws of thermodynamics is that in many situations they can be used to predict the maximum efficiency that could be achieved by a perfect machine, without specifying any details of the machine! (Efficiency may be defined, in this situation, as the ratio of useful work to total energy flow.) Thus, one can specify, for example, what *minimum* amount of energy is necessary to separate salt from seawater, to separate metals from their ores, and to separate pollutants from auto exhaust without knowing any details about future inventions that might be devised for these purposes. Similarly, if one is told the temperature of a source of thermal energy—say, the hot rock deep in Earth's crust—one can calculate rather easily the maximum efficiency with which this thermal energy can be converted to applied work, regardless of the cleverness of future inventors. In other words, *there are some fixed limits to technological innovation, placed there by fundamental laws of nature.* (The question of how far from the maximum attainable efficiencies industrial societies operate today is taken up in Chapter 8.)

More generally, the laws of thermodynamics explain why we need a continual input of energy to maintain ourselves, why we must eat much more than a pound of food in order to gain a pound of weight, and why the total energy flow through plants will always be much greater than that through plant-eaters, which in turn will always be much greater than that through flesh-eaters. They also make it clear that *all* the energy used on the face of the Earth, whether of solar or nuclear origin, will ultimately be degraded to heat. Here the laws catch us both coming and going, for they put limits on the efficiency with which we can manipulate this heat. Hence, they pose the danger (discussed further in Chapter 11) that human society may make this planet uncomfortably warm with degraded energy long before it runs out of high-grade energy to consume.

Occasionally it is suggested erroneously that the process of biological evolution represents a violation of the second law of thermodynamics. After all, the development of complicated living organisms from primordial chemical precursors, and the growing structure and complexity of the biosphere over the eons, do appear to be the sort of spontaneous increases in order excluded by the second law. The catch is that Earth is not an isolated system; the process of evolution has been powered by the sun, and the decrease in entropy on Earth represented by the growing structure of the biosphere is more than counterbalanced by the increase in the entropy of the sun. (The process of evolution is discussed in more detail in Chapter 4.)

It is often asked whether a revolutionary development in physics, such as Einstein's theory of relativity, might not open the way to circumvention of the laws of thermodynamics. Perhaps it would be imprudent to declare that in no distant corner of the universe or hitherto-unexplored compartment of subatomic matter will any exception ever turn up, even though our intrepid astrophysicists and particle physicists have not yet found a single one. But to wait for the laws of thermodynamics to be overturned as descriptions of everyday experiences on this planet is, literally, to wait for the day when beer refrigerates itself in hot weather and squashed cats on the freeway spontaneously reassemble themselves and trot away.

fraction by volume, often used elsewhere for gases. In many applications it is also useful to work with the *mass fraction* (grams of constituent/gram of mixture). This is a more precise statement of what is meant by the common term *fraction by weight* or *percent by weight.*[21]

A *mole* of any substance is 6.02×10^{23} molecules, and the mass of a mole is equal to the molecular weight of the substance in grams. For example, the mass of a mole of nitrogen gas (N_2) is 28 grams. It is often convenient to speak of air as if it were a single substance; the term *a mole of air* means 6.02×10^{23} molecules, of which 78.09 percent are nitrogen molecules, 20.95 percent are oxygen molecules, and so on. Such a collection of molecules has a mass of about 29 grams, which is called the *molecular weight of air.* (These definitions will be of importance later in interpreting what is meant by pollution standards expressed in different ways.)

Although nitrogen and oxygen comprise 99 percent of dry air, the trace constituents carbon dioxide (CO_2) and ozone (O_3) play exceedingly important roles because of the special properties of these molecules, as described below. Methane, nitrous oxide, and hydrogen also have roles in atmospheric chemistry and physics, albeit smaller ones. Argon, helium, krypton, and neon, by contrast, are chemically inert, monatomic gases, whose presence in the atmosphere is of interest only as resources for certain applications in technology.[22]

Water Vapor

The water content of the atmosphere varies greatly from place to place and time to time. Three commonly used measures of water content are *absolute humidity, specific humidity,* and *relative humidity.* Absolute humidity is the mass of water vapor per unit volume of air, and it varies from almost zero over the driest deserts to around 25 grams per cubic meter over jungles and tropical seas. Specific humidity is the mass of water vapor per unit mass of air. (A closely related term, the *mixing ratio,* is the mass of water vapor mixed with each unit mass of *dry* air.) Relative humidity, usually expressed as a percentage, is the ratio of the actual molecular fraction of water vapor in air to the molecular fraction corresponding to saturation at the prevailing temperature. (*Saturation* refers to the condition that ensues if air is left for a long time in a sealed container partly filled with pure water; the number of molecules of water vapor per unit volume of air under these circumstances depends only on the temperature.)[23] Relative humidity usually is between 0 and 100 percent, but under special circumstances (supersaturation) it can significantly exceed 100 percent.

Under ordinary circumstances, water vapor in the atmosphere begins to condense into droplets of liquid, forming clouds, as soon as the relative humidity exceeds 100 percent by even a small amount. The process of condensation is greatly facilitated by the virtually universal presence in the atmosphere of small particles that provide surfaces where the condensation can commence. Called *condensation nuclei* when they perform this function, these particles include salt crystals formed by the evaporation of sea spray, dust raised by the wind, ash from volcanoes and forest fires, decomposed organic matter, and, of course, particles produced by various technological activities. Even in "unpolluted" air, particles that might serve as condensation nuclei are seemingly abundant in absolute terms (more than 100 particles per cubic centimeter), but the extent of condensation and precipitation apparently are related to specific physical characteristics of the condensation nuclei as well as to their number.

The molecular fraction of water vapor corresponding to saturation increases as temperature increases—warm air can "hold" more water vapor than cool air. Accordingly, there are two ways in which relative humidity can be raised from less than 100 percent to more, initiating

---

[21] Weight means the force exerted upon a mass by gravity. Weight and mass are more or less interchangeable (using the relation, weight equals mass multiplied by acceleration of gravity) only if one stays on Earth's surface, where gravity is nearly constant. An astronaut has the same mass on the surface of the moon as on Earth, but a very different weight, because the acceleration of gravity on the moon is much less than on Earth.

[22] For discussions of how the atmosphere came to have the composition it does, the reader should consult Preston Cloud and Aharon Gibor, The oxygen cycle; and Preston Cloud, ed., *Adventures in earth history.*

[23] The often encountered definition of saturation as "the maximum amount of water vapor air can hold at a given temperature" is not quite correct. A good discussion of this and the following points is given by Morris Neiburger, James G. Edinger, William D. Bonner, *Understanding our atmospheric environment,* Chapter 8.

condensation and perhaps precipitation: (1) more water vapor can be added to the air by evaporation from an exposed water surface; (2) the air can be cooled so that the vapor content corresponding to saturation falls. At a given vapor content (a fixed specific humidity), the temperature at which the relative humidity reaches 100 percent is called the *dew point.* Addition of water vapor to the air by evaporation is a slow process, but cooling of the air can be very rapid. Rapid cooling to below the dew point is the mechanism immediately responsible for most condensation phenomena – the appearance of dew and fog at night as air is cooled by radiation of heat to the night sky; formation of clouds and rain in updrafts as the air is cooled by expansion; and formation of beads of water on the outside of a pitcher of ice water on a hot day, as air adjacent to the pitcher is cooled by contact with the cold surface.

Pressure, Temperature, and Vertical Structure

The pressure exerted by the atmosphere on objects at Earth's surface is essentially equal to the weight of the overlying air, which at sea level amounts on the average to 10.3 metric tons per square meter (14.7 lb/in^2). This amount of pressure, defined as 1 atmosphere, is the same would be exerted at sea level by a column of water about 10 meters high (33 ft) or a column of mercury 760 millimeters (mm) high (29.92 in). This means that the mass of the atmosphere is only equivalent to that of a 10-meter layer of water covering Earth (you can check this in Table 2-2) and that pressure under water increases by the equivalent of 1 atmosphere for every 10 meters of depth. The usual metric unit for the measurement of atmospheric pressure is the millibar; 1 millibar is 100 newtons (N) per square meter (see Box 2-1 for the definition of the newton), and 1 atmosphere is 1013.25 millibars.

Atmospheric pressure is not ordinarily perceived as a force because it acts equally in all directions (up, down, sideways); organisms are not crushed by it because the gases and liquids in tissue are also at atmospheric pressure, so the inward and outward forces balance. Pressure becomes perceptible as a (painful) pressure *difference* if the pressure outside an organism changes more rapidly than the interior pressure can accommodate (an example would be the pain in one's ears associated with a rapid change in altitude).

Unlike water, whose density at the bottom of the deepest ocean trenches at pressures of hundreds of atmospheres is only a few percentage points higher than its density at the surface, the air in the atmosphere is highly compressible – that is, density increases markedly as pressure increases. Indeed, air behaves very much like a "perfect gas," for which pressure (p), density (ρ), and temperature (T) are related by the equation

$$p = \rho R T$$

For this equation to be valid, the temperature T must be measured with respect to ***absolute zero,*** the temperature at which there is no molecular motion. Temperature measured from this zero point, which is the same for all substances, is called ***absolute temperature,*** and the corresponding unit of measurement in the metric system is the degree kelvin.[24] The R in the equation is the gas constant, which for dry air equals 287 joules per kilogram per degree kelvin. According to the perfect gas equation, the density of dry air varies in direct proportion to pressure, if temperature is held constant.

Because the atmosphere is compressible, its mass is concentrated in the lower layers. Forty percent of the air in the atmosphere lies below the altitude of the summit of Mount Whitney in California's Sierra Nevada range (4.4 km) and two-thirds lies below the altitude of the summit of Mount Everest (8.9 km). The density of air at an altitude of 12 kilometers, where most subsonic jet airliners fly, is about one-fifth the density at sea level. The average variation of pressure and temperature with altitude above sea level is shown in Figure 2-10.

The atmosphere is subdivided into horizontal layers according to the pattern of temperature variation. The lowest layer, called the ***troposphere,*** is characterized by a rather uniform average rate of temperature decline with altitude of 6.4° C per kilometer. Almost all the atmo-

---

[24] Absolute zero, or 0 degrees kelvin (K) equals −273.15° C. An attempt was made recently to standardize the unit of absolute temperature as simply the kelvin, rather than the degree kelvin, but the change has not been generally adopted.