

Thermodynamics

The science of thermodynamics originally evolved from a variety of observational evidence indicating that mechanical energy is transformable into heat. For example, Benjamin Thompson (1753–1814) of Woburn, Massachusetts, better known as Count Rumford of the Holy Roman Empire, was impressed by the amount of heat generated during the boring of cannon at the arsenal in Munich. In 1798 he proposed that the heat was formed by conversion of the mechanical energy expended by the boring and actually calculated the amount of heat generated by one horse turning the borer for one hour. Count Rumford's ideas about the equivalence of work and heat were supported by Sir Humphry Davy (1778–1829), who demonstrated in 1799 that two pieces of ice could be made to melt by rubbing them together by a clockwork mechanism in a vacuum. Nevertheless, the scientific community did not accept this idea for nearly 50 years until James P. Joule (1818–1889), working in his father's brewery in Manchester, England, carefully measured the rise in temperature caused by stirring water with a mechanical paddle wheel. His research between 1840 and 1849 clearly demonstrated that the expenditure of a certain amount of work always produced the same amount of heat (Moore, 1955).

The observation that work is transformable into heat is an extension of the principle of conservation of energy. It can be restated by saying that whenever work is done, heat energy is consumed, implying that one must always pay a price by the expenditure of energy if work is to be accomplished.

Thermodynamics is founded on three basic principles, which arise from common experience

but cannot be proved or disproved. They must be accepted as articles of faith. The science of thermodynamics has evolved from these so-called laws by careful deductive reasoning aided by mathematics.

11.1 Definitions

Thermodynamics describes the world the way it appears to an observer in terms of certain measurable properties of matter, such as volume, pressure, temperature, and chemical composition. Matter is regarded as continuous and indivisible. In this regard thermodynamics is the opposite of *statistical mechanics* in which matter is treated as a collection of particles in motion. In order to enter the world of thermodynamics we must first become familiar with certain definitions and conventions that may differ from common usage.

The basic entity for consideration in thermodynamics is the *system*, which is that part of the universe whose properties are under consideration. The system is separated from the rest of the universe, known as the *surroundings*, by a *boundary* whose properties can be defined. A system is said to be *open* when matter can pass across the boundary and *closed* when matter is prevented from entering or leaving the system. The boundary may prevent not only matter but also heat and any other form of energy from passing through it. Systems that have such impervious boundaries are said to be *isolated*. The investigator must identify the system to be studied and specify the properties of its boundary.

The properties of a system are classified as being either *intensive* or *extensive*. Intensive properties, such as pressure, temperature, and

density, are independent of the amount of matter that is present, whereas the extensive properties, such as volume and mass, depend on the mass and are therefore additive. When the properties of a system are changed, it experiences a change in its *state*. For example, when a quantity of gas in a cylinder is compressed by moving a frictionless piston, the system undergoes a *change in state*.

The chemical composition of thermodynamic systems is expressed in terms of *components* and *phases*. Nordstrom and Munoz (1986, p. 67) defined a *phase* as "a uniform, homogeneous, physically distinct, and mechanically separable portion of a system." When the system under consideration is a rock, the minerals of which the rock is composed are the phases. Phases need not be solids, but can also be liquids or gases. For example, a system containing ice, water, and water vapor has three phases (solid, liquid, and gas). The aqueous phase in such a system may also contain a number of ionic or molecular *species* in solution.

The *components* of a system are the chemical constituents by means of which the chemical compositions of the phases of a system can be completely described. The choice of components is arbitrary. For example, the chemical composition of a rock specimen can be expressed in terms of oxides or in terms of chemical elements. However, the *number* of components required to describe the compositions of all of the phases present in a system at equilibrium is constrained by the *Gibbs phase rule*.

The number of variables that must be specified in order to define the state of a system is equal to the number of components plus the number of intensive properties, such as pressure and temperature. The state of the system is completely defined when the number of variables is equal to the number of independent relations between them. If the number of independent equations is *less* than the number of variables, the system has *variance* or *degrees of freedom*. In this case, the state of the system cannot be completely constrained. When the number of variables *exceeds* the number of independent equations by one, the system is *univariant* and has one degree of freedom. Univariant systems can be con-

strained by specifying the value of *one* of the variables. When the number of variables exceeds the number of equations by *two*, the system is *divariant*, and so on.

When a chemical system is at equilibrium, the number of phases (p), the number of components (c), and the number of degrees of freedom (f) must satisfy the Gibbs phase rule (Nordstrom and Munoz, 1986):

$$p + f = c + 2 \quad (11.1)$$

Therefore, systems in which the number of phases is *equal* to the number of components have two degrees of freedom. In other words, pressure and temperature must be specified in this case in order to completely define the state of the system.

All systems contain a certain amount of *heat*, which is defined as that which flows across the boundary of a closed system during a change in its state by virtue of a difference in temperature between the system and its surroundings. Moreover, heat always flows from a point of higher to a point of lower temperature. Thermodynamics defines heat as though it were an invisible fluid whose *direction* of flow is indicated by its algebraic sign. Heat is *positive* when it flows from the surroundings to the system and *negative* when it flows from the system to the surroundings. We can remember this important convention by recognizing that systems are "selfish." When a system gains heat, the added heat has a positive sign, and when it loses heat the sign is negative.

Work is defined in mechanics as the product of force times distance, whereas in thermodynamics work is defined as a quantity that flows across the boundary of a system during a change in its state and is completely convertible into the lifting of a weight in the surroundings. Work, like heat, is an algebraic quantity, but the sign convention of work is opposite to that of heat. When work flows to the surroundings it has a positive sign, and when it flows to the system it has a negative sign. Another way to visualize the sign convention for work is to say that when the system does work, the work is positive, but when the surroundings do it, the work is negative.

With these definitions and conventions we have now set the stage to introduce the first law of

thermodynamics. The presentation will be selective and based on plausibility rather than mathematical rigor. More complete and rigorous presentations can be found in college level textbooks on physical chemistry. Thermodynamics for use by geologists is presented in books by Kern and Weisbrod (1967), Wood and Fraser (1976), Fraser (1977), Nordstrom and Munoz (1986), and Anderson (1996).

11.2 The First Law

Every system whose state has been defined by specifying the numerical values of its properties contains a certain amount of energy. This energy arises from the heat content of the system as well as from other internal sources such as gravitational, electrical, magnetic, molecular, atomic, and even nuclear phenomena. The amount of internal energy (E) of a particular system is difficult to determine. However, its magnitude depends on the state of the system. When a system undergoes a change in its state in a reversible manner, its internal energy also changes. The magnitude of the change in the internal energy ΔE of a system that undergoes a reversible change in state from state 1 to state 2 is given by:

$$\Delta E = E_2 - E_1 \quad (11.2)$$

where E_2 and E_1 are the internal energies of the system in states 2 and 1, respectively. It makes no difference *how* the change in state is carried out, provided that it is *reversible*, which requires that the system is never far from being in equilibrium. If the system is returned to its original state, the net change in its internal energy is zero. We conclude therefore that the change in the internal energy of a system undergoing a change in state along a circular path is zero:

$$\Delta E = (E_2 - E_1) + (E_1 - E_2) = 0 \quad (11.3)$$

The *first law of thermodynamics* states that the increase in the internal energy of a system during a reversible change in its state is equal to the heat that flows across its boundary from the surroundings (q) minus the work done by the system (w).

Therefore, the net change in the internal energy of the system is:

$$\Delta E = q - w \quad (11.4)$$

The first law of thermodynamics implies that work and heat are related and that some of the heat that flows into the system during a change in state is converted into work and therefore does not augment the internal energy.

We can imagine that a reversible change in state can be carried out in small incremental steps such that the internal energy changes by small amounts designed by dE . Now we can restate the first law in differential form:

$$dE = dq - dw \quad (11.5)$$

where dq and dw are very small increments of heat and work, respectively. The most common form of work a system can do during a change in its state is to expand against the constant pressure of the surroundings. Thus we can specify that:

$$dw = PdV \quad (11.6)$$

where P is the external pressure and dV is an incremental change in the volume of the system. A simple demonstration indicates that PdV does have the dimensions of work (force $\times$ distance):

$$PdV = \frac{\text{force} \times (\text{distance})^3}{(\text{distance})^2} = \text{force} \times \text{distance} \quad (11.7)$$

Therefore, we can express the first law by the equation:

$$dE = dq - PdV \quad (11.8)$$

11.3 Enthalpy

The statement of the first law represented by equation 11.8 enables us to predict the change in the internal energy of a system during a change in its state. The equation indicates that the increase in the internal energy of the system is equal to the heat absorbed from the surroundings minus the work done by the expansion of the system against the pressure exerted by the surroundings. We can

carry out such a change in state by integrating equation 11.8 from state 1 to state 2:

$$\int_1^2 dE = \int_1^2 dq - P \int_1^2 dV \quad (11.9)$$

The integration yields:

$$E_2 - E_1 = (q_2 - q_1) - P(V_2 - V_1) \quad (11.10)$$

We simplify this result by replacing $q_2 - q_1$ by q_P , which is an *amount* of heat added at constant pressure. In addition, we multiply the PV terms:

$$E_2 - E_1 = q_P - PV_2 + PV_1 \quad (11.11)$$

By rearranging the terms of this equation we obtain:

$$(E_2 + PV_2) - (E_1 + PV_1) = q_P \quad (11.12)$$

It is advantageous at this point to define a new kind of energy called the *enthalpy* (H):

$$H = E + PV \quad (11.13)$$

The enthalpy is a function of the state of the system, as is the internal energy. Therefore, the enthalpy of the system in states 1 and 2 is:

$$H_1 = E_1 + PV_1 \quad (11.14)$$

$$H_2 = E_2 + PV_2 \quad (11.15)$$

Substituting equations 11.14 and 11.15 into equation 11.12 yields:

$$H_2 - H_1 = q_P = \Delta H \quad (11.16)$$

In other words, the change in the enthalpy of a system (ΔH) during a reversible change in its state at constant pressure is equal to the heat absorbed by the system during that change in state.

This is a very important and concrete result because q_P is a *measurable* quantity. For example, when a chemical reaction is carried out in a well-insulated vessel called a calorimeter, the heat evolved or consumed is indicated by the change in temperature inside the calorimeter. Therefore, equation 11.16 enables us to measure the *change* in the enthalpy of a system as a result of a chemical reaction, but we still do not know the *amount* of enthalpy the system contains.

In order to establish a *scale* for enthalpy we now define a *reference state* for chemical elements and

their compounds. The reference state is defined in terms of certain *standard properties* that define the so-called *standard state*. The temperature of the standard state is equal to 25°C (298.15 K), the pressure is 1 atm (recently redefined as 1 bar, which is equal to 10^6 dyne/cm², or 0.987 atm, 29.53 in. Hg), and the activity of ions and molecules in aqueous solution is equal to one. In addition, we specify that the enthalpies of the pure elements in their stable state of aggregation in the standard state are equal to zero. Thus we have the definition:

The *enthalpy of formation* of compounds and their ions and molecules in aqueous solution is the heat absorbed or given off by chemical reactions in which the compounds, ions, and molecules form from the elements in the standard state.

We represent this important thermodynamic parameter by the symbol H_f° , where the superscript $^\circ$ identifies the standard state and the subscript f symbolizes "formation."

The standard enthalpies of formation and other thermodynamic constants of many compounds, ions, and molecules have been measured experimentally and are tabulated in the *CRC Handbook of Chemistry and Physics* (Weast et al., 1986), in Technical Notes of the National Bureau of Standards, and in compilations by Kelley (1962), Robie et al. (1978), Barin et al. (1989), and Woods and Garrels (1987). However, it is important to remember that standard enthalpies of formation and other thermodynamic constants are *measurements* and are subject to experimental errors. Consequently, the values reported by different investigators for a given compound commonly differ from each other by amounts that may or may not be within the error of the measurements. Such discrepancies in thermodynamic constants affect the results of calculations in which they are used and cause internal inconsistencies. Some of the standard enthalpies of formation listed in Appendix B are average values of two or more determinations that appear to be compatible within the errors of measurement. However, no assurance can be given that these values are accurate to within the stated number of significant figures or that the listed values are internally consistent.

11.4 Heats of Reaction

When two compounds or elements, A and B, react to form a product A_2B_3 by the reaction:



a certain amount of heat is either used up or given off. We can now calculate the amount of this heat (q_p) from the enthalpy change of the system when the reaction takes place in the standard state. In other words, we treat the reaction as a change in state and calculate the difference in the enthalpy (ΔH_R°) between the final state (A_2B_3) and the initial state ($2A + 3B$). Therefore, the change in enthalpy of the reaction in the standard state is:

$$\Delta H_R^\circ = H_f^\circ(A_2B_3) - [2H_f^\circ(A) + 3H_f^\circ(B)] \quad (11.18)$$

where ΔH_R° is the heat of the reaction in the standard state and H_f° is the standard enthalpy of formation of A_2B_3 , A, and B, respectively. Note that the heat is evolved or absorbed as a result of, or during, the reaction and *not* when the reaction is at equilibrium. After a reaction has reached equilibrium, the heat of the reaction (ΔH_R) is zero because the rates of the forward and backward reactions are exactly equal. Thus, we can state that at equilibrium $\Delta H_R = 0$. However, the reference state is a defined condition in which equilibrium does not occur. Therefore ΔH_R° is never zero. In general, ΔH_R° is calculated by summing the standard enthalpies of the products and by subtracting the enthalpies of the reactants:

$$\Delta H_R^\circ = \sum_i n_i H_{fi}^\circ(\text{products}) - \sum_i n_i H_{fi}^\circ(\text{reactants}) \quad (11.19)$$

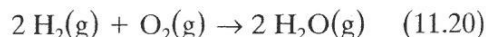
where n is the molar coefficient of each reactant and product taken from a balanced equation representing the reaction and i identifies the compounds or ions that participate in the reaction.

Standard enthalpies are expressed in units of kilocalorie per mole (kcal/mol), where the normal *calorie* is defined as the amount of heat required to raise the temperature of 1 g of water from 14.5 to 15.5°C. The calorie is equivalent to

4.1840 joules (J) so that 1 kcal/mol is equal to 4.1840 kJ/mol. At the present time both units are in use.

When ΔH_R° of a reaction is positive, the reaction is *endothermic*, which means that heat flows from the surroundings to the system. Conversely, when ΔH_R° is negative, the reaction is *exothermic* and heat flows from the system to the surroundings. Endothermic reactions consume heat, whereas exothermic reactions produce heat.

In order to illustrate how heats of reaction are calculated, we choose the reaction between hydrogen and oxygen to form water in the gaseous state, symbolized by (g):



The amount of heat absorbed or given off when this reaction occurs at 1 atm and 25°C is given by:

$$\Delta H_R^\circ = 2H_f^\circ(H_2O(g)) - [2H_f^\circ(H_2(g)) + H_f^\circ(O_2(g))] \quad (11.21)$$

The standard enthalpies of H_2 and O_2 are both equal to zero by definition, so that in this case:

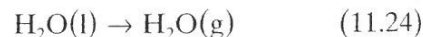
$$\Delta H_R^\circ = 2H_f^\circ(H_2O(g)) \quad (11.22)$$

The standard enthalpy of formation of water vapor is -57.80 kcal/mol (Appendix B). Therefore:

$$\Delta H_R^\circ = 2(-57.80) = -115.6 \text{ kcal} \quad (11.23)$$

We conclude that this reaction is strongly *exothermic*. In fact, this reaction is commonly used to demonstrate the explosive nature of some chemical reactions. Note that the reaction is *endothermic* when equation 11.20 is *reversed*.

Heat effects are also associated with the dissociation of salts into ions, with phase transformations, and with mixing or dilution of solutions. For example, heat is absorbed from the surroundings when water evaporates to form water vapor:



When the evaporation occurs in the standard state:

$$\Delta H_R^\circ = H_f^\circ(H_2O(g)) - H_f^\circ(H_2O(l)) \quad (11.25)$$

From Appendix B:

$$\Delta H_R^\circ = (-57.80) - (-68.32) = +10.52 \text{ kcal} \quad (11.26)$$

Evidently, the evaporation of water at 25°C and 1 atm pressure is an *endothermic* process that removes heat from the surroundings and thereby causes the temperature to decrease. The heat required to evaporate one mole of a liquid is called the *latent heat of evaporation*. When the vapor condenses to form the liquid, the heat given off is called the *latent heat of condensation*. In other words, evaporation is endothermic, whereas condensation of water vapor is exothermic, and the latent heat of evaporation is equal in magnitude to the heat of condensation. Similarly, fusion or melting of solids is endothermic, whereas crystallization of liquids is exothermic, and the *latent heat of fusion* is equal to the *latent heat of crystallization*.

Note that we are unable to calculate the latent heat of fusion or crystallization of water because it takes place at 0°C rather than at 25°C. Therefore, in order to calculate the latent heat of fusion of water, we must first extrapolate the enthalpies of formation from 25 to 0°C. In general, chemical reactions on the surface of the Earth occur at temperatures between 0 and 100°C, and their heats of reaction therefore commonly differ from those calculated at 25°C.

The response of a reaction *at equilibrium* to a change in temperature can be predicted on the basis of Le Châtelier's principle. If the forward reaction is exothermic, an *increase* in the temperature favors the backward reaction because it consumes heat and therefore counteracts the increase in the temperature. Consequently, the equilibrium will shift to the left making the equilibrium constant smaller. Similarly, a *decrease* in temperature applied to a reaction at equilibrium in which the forward reaction is exothermic causes the equilibrium to shift to the right and increases the value of the equilibrium constant. Evidently, the enthalpy change of a chemical reaction is related to the way in which the numerical value of the equilibrium constant varies with temperature. This relationship is formally expressed by the van't Hoff equation, which will be derived in Section 11.10.

We may also expect that exothermic reactions take place spontaneously because they do not require heat from the surroundings. However, this generalization is not correct. It turns out that not

all of the heat that enters a system during a reversible change in its state is available to do work. Similarly, the direction in which a reaction proceeds spontaneously is not determined solely by the enthalpy change because systems have an additional property called *entropy*. We return to this phenomenon in Section 11.6.

11.5 Heat Capacity

When heat is added to a solid, a liquid, or a gas, the temperature of the substance increases. This familiar observation can be stated as an equation:

$$dq = C dT \quad (11.27)$$

where dq is an increment of heat added, C is a constant known as the *heat capacity*, and dT is the corresponding incremental increase in the temperature expressed in *kelvins* (K). In thermodynamics, temperatures are always expressed on the "absolute" or Kelvin scale, which has its zero point at -273.15°C , when all molecular motion stops and the heat content of pure crystalline materials becomes equal to zero.

The heat capacity is a characteristic property of chemical elements and their compounds that must be measured experimentally. It turns out that heat capacities actually vary with temperature and that it makes a difference whether the heat is added at constant pressure or at constant volume. On the surface of the Earth pressure is constant, and we therefore work with the "heat capacity at constant pressure" symbolized by C_p . Consequently, we restate equation 11.27 as:

$$dq_p = C_p dT \quad (11.28)$$

We now recall that the heat added to a closed system during a reversible change in its state at constant pressure is equal to the increase in the enthalpy of the system. This conclusion was expressed by equation 11.16, which we now restate in terms of differentials:

$$dq_p = dH \quad (11.29)$$

Substituting equation 11.29 into 11.28 yields:

$$dH = C_p dT \quad (11.30)$$

This equation can be integrated between appropriate temperature limits and therefore enables us to extrapolate enthalpies of formation of compounds from the standard temperature (T°) to some other temperature (T) above or below it:

$$\int_{T^\circ}^T dH = C_p \int_{T^\circ}^T dT \quad (11.31)$$

By carrying out the integration we obtain:

$$H_T - H_{T^\circ} = C_p(T - T^\circ) \quad (11.32)$$

Equation 11.32 enables us to calculate the enthalpy of formation of a compound at temperature T if we know its standard enthalpy and its heat capacity, *provided* the temperature dependence of C_p can be neglected.

In reality, use of equation 11.32 to extrapolate enthalpies is limited because the heat capacity is in fact a function of the temperature and therefore cannot be treated as a constant in the integration. Moreover, the range of temperatures over which the integration is carried out must not include phase transformations, such as boiling or freezing, which cause a discontinuity in the heat capacity and involve their own enthalpy changes. These restrictions generally are not severe provided the environmental conditions remain between 0 and 100°C at 1 atm pressure. Therefore, the most straightforward procedure is to recalculate the enthalpies of formation of reactants and products to the desired temperature (between 0 and 100°C) and then to calculate the enthalpy change of the reaction at that temperature.

A more accurate and elegant procedure is based on equations that express the variation of C_p as a function of temperature in the form:

$$C_p = a + (b \times 10^{-3})T + (c \times 10^{-6})T^2 \quad (11.33)$$

where a , b , and c are constants derived by fitting an algebraic equation to experimentally determined data in coordinates of C_p and T . Substituting equation (11.33) into equation 11.31 yields:

$$\int_{T^\circ}^T dH = a \int_{T^\circ}^T dT + b \times 10^{-3} \int_{T^\circ}^T T dT$$

$$+ c \times 10^{-6} \int_{T^\circ}^T T^2 dT \quad (11.34)$$

By integrating this equation we obtain:

$$H_T - H_{T^\circ} = a(T - T^\circ) + \frac{b \times 10^{-3}}{2} [T^2 - (T^\circ)^2] + \frac{c \times 10^{-6}}{3} [T^3 - (T^\circ)^3] \quad (11.35)$$

Equation 11.35 can be applied to the reactants and products of a chemical reaction as specified by equation 11.19. As a result, we calculate Δa , Δb , and Δc for the reaction and then evaluate the enthalpy change of the reaction at temperature T ($\Delta H_{R,T}$) from the equation:

$$\Delta H_{R,T} - \Delta H_R^\circ = \Delta a(T - T^\circ) \quad (11.36)$$

$$+ \frac{\Delta b \times 10^{-3}}{2} [T^2 - (T^\circ)^2] + \frac{\Delta c \times 10^{-6}}{3} [T^3 - (T^\circ)^3]$$

The relationship between C_p and T is sometimes expressed in ways that differ from equation 11.33, but the application of these equations is limited by the availability of experimental data.

11.6 The Second Law

In the 19th century much work was done by means of steam engines in which hot steam expands in a cylinder and pushes a piston. The steam cools as it expands and is returned to the reservoir at a lower temperature. An important scientific problem of that time was to understand the factors that limit the *efficiency* of steam engines. This problem was solved in 1824 by the French engineer Sadi Carnot (1796–1832). Without going into the details, we can say that the efficiency of steam engines can never be 100% because some of the heat added to the system must be discharged when the expanded steam returns to the reservoir.

The work of Carnot was a very important contribution to thermodynamics and from it arose a new principle known as the *second law of thermodynamics*. The essence of the second law is

that all systems possess a property called *entropy*, named by the German physicist Rudolf Clausius (1822–1888). One way to explain entropy is to say that a certain fraction of the enthalpy of a system is not convertible into work because it is consumed by an increase in the entropy. Another way to explain the meaning of entropy is to compare it to the property of “randomness” recognized in *statistical mechanics*. When one tries to predict the behavior of a collection of particles, one comes to the conclusion that they will most probably assume a condition of maximum randomness. This conclusion leads to the generalization that every system that is left to itself will, on the average, change toward a condition of maximum randomness. The implication of this statement is that the entropy of a system increases spontaneously and that energy must be spent to reverse this tendency.

When entropy is thought of as randomness, it can be recognized in many natural phenomena. We have all learned from the common experiences of life that work must be done to counteract the natural tendency toward disorder. Many popular science books have been written about entropy and its manifestations in everyday life including one by Rifkin (1980) who claimed that “entropy is the supreme law of nature and governs everything we do.”

The second law introduces entropy by the statement:

In any *reversible* process the change in the entropy of the system (dS) is equal to the heat received by the system (dq) divided by the absolute temperature (T):

$$dS = \frac{dq}{T} \quad (\text{reversible process}) \quad (11.37)$$

In any spontaneous *irreversible* process the change in entropy is greater than this amount:

$$dS > \frac{dq}{T} \quad (\text{irreversible process}) \quad (11.38)$$

Entropy, like internal energy and enthalpy, is a property whose magnitude depends on the state

of the system. However, unlike internal energy and enthalpy, the absolute value of the entropy of a system can be determined by virtue of the third law of thermodynamics.

Measurements indicate that the heat capacities of pure crystalline materials *decrease* with decreasing temperature and approach zero at absolute zero. This evidence has been elevated into the *third law of thermodynamics*, which states:

The heat capacities of pure crystalline substances become zero at absolute zero.

Since $dq = CdT$ (equation 11.28) and $dS = dq/T$ (equation 11.37), it follows that:

$$dS = C \left(\frac{dT}{T} \right) \quad (11.39)$$

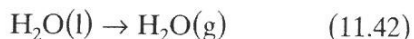
If the heat capacity is equal to zero at absolute zero, then $dS = 0$. Therefore, when a system composed of pure crystalline substances undergoes a reversible change in state at absolute zero, the entropy change ΔS is equal to zero. The German physicist Walther Hermann Nernst (1864–1941) expressed these insights by stating that at absolute zero the molar entropies of pure crystalline solids are equal to zero. Therefore, the third law enables us to calculate entropies by integrating equation 11.39 from absolute zero to some higher temperature (T):

$$\int_0^T dS = \int_0^T C \left(\frac{dT}{T} \right) \quad (11.40)$$

In order to calculate the entropy of an element or compound in the standard state from equation 11.40, the temperature dependence of its heat capacity must be determined experimentally. The standard molar entropies of many elements and compounds have been determined and are listed in reference books including the *CRC Handbook of Chemistry and Physics* (Weast et al., 1986). Molar entropies are expressed in “entropy units” or calories per degree (cal/deg). When the standard molar entropies of the reactants and products of a chemical reaction are known, the change in the entropy (ΔS_R°) of a chemical reaction or phase transformation can be calculated from the statement:

$$\Delta S_R^\circ = \sum n_i S_i^\circ(\text{products}) - \sum n_i S_i^\circ(\text{reactants}) \quad (11.41)$$

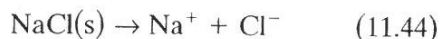
When ΔS_R° is positive, the entropy of the system increases as a result of the change in state, whereas when ΔS_R° is negative, it decreases. For example, when liquid water evaporates in the standard state to form water vapor:



the change in entropy is (Krauskopf, 1979):

$$\Delta S_R^\circ = 45.10 - 16.71 = +28.39 \text{ cal/deg} \quad (11.43)$$

The increase in the entropy is consistent with our expectation that water molecules in the gas phase are more randomly distributed than molecules of liquid water. Similarly, when solid NaCl dissociates to form ions:



the change in entropy is (Krauskopf, 1979):

$$\Delta S_R^\circ = 14.0 + 13.5 - 17.2 = +10.3 \text{ cal/deg} \quad (11.45)$$

Evidently, the entropy increases when NaCl dissolves to form ions in aqueous solution because the ions occupy fixed lattice positions in the crystal whereas the ions in the solution are mobile. However, for other reactions ΔS_R° may be negative, which implies a *reduction* in the entropy of the system.

11.7 Gibbs Free Energy

The second law of thermodynamics implies that the *increase* in the *enthalpy* of a system during a reversible change in its state at constant temperature is diminished because a certain amount of the enthalpy is consumed by an increase in the *entropy* of the system. This insight was used by J. Willard Gibbs (1839–1903) to define a new form of energy, now known as the *Gibbs free energy* (G):

$$G = H - TS \quad (11.46)$$

where H is the enthalpy and S is the entropy. The change in the Gibbs free energy of a chemical reaction in the standard state is:

$$\Delta G_R^\circ = \Delta H_R^\circ - T\Delta S_R^\circ \quad (11.47)$$

The Gibbs free energy is a function of the state of the system like the internal energy and the enthalpy. It is therefore subject to the same conventions regarding the standard state as enthalpy and is measured in kilocalories per mole:

The *standard Gibbs free energy of formation* of a compound (G_f°) is the change in the free energy of the reaction by which it forms from the elements in the standard state.

Numerical values of G_f° for compounds and ions in aqueous solution are listed in Appendix B together with values of H_f° . The value of ΔG_R° for a reaction is calculated by summing the values of G_f° of the products, each multiplied by its molar coefficient, and by subtracting from it the sum of the G_f° values of the reactants:

$$\Delta G_R^\circ = \sum n_i G_{fi}^\circ(\text{products}) - \sum n_i G_{fi}^\circ(\text{reactants}) \quad (11.48)$$

The algebraic sign and magnitude of ΔG_R° depend on the sign and magnitude of ΔH_R° and ΔS_R° in equation 11.47. If the forward reaction is exothermic (ΔH_R° is negative) and the entropy increases (ΔS_R° is positive), then the two terms of equation 11.47 combine to make ΔG_R° negative. If, however, the entropy decreases (ΔS_R° is negative), then $T\Delta S_R^\circ$ in equation 11.47 is positive. If ΔH_R° is negative, ΔG_R° will be less negative than ΔH_R° and may even become positive if $T\Delta S > \Delta H$. Note that $T\Delta S$ is in fact a form of enthalpy because $\Delta S = \Delta H/T$ at constant pressure and temperature (equations 11.16 and 11.37). Therefore, $T\Delta S$ has the dimensions of ΔH and the two terms can be added or subtracted from each other.

When ΔG_R° of a reaction is *negative*, the forward reaction has excess energy when it occurs in the standard state. When ΔG_R° is *positive*, there is

a deficiency of energy for the forward reaction in the standard state and the reaction will therefore run in the backward direction. In general, ΔG_R° is a measure of the *driving force* of a chemical reaction such as equation 11.17:



If ΔG_R° of this reaction is negative, the reaction proceeds as written from left to right in the standard state. When the reaction ultimately reaches equilibrium, the product is more abundant than the reactants so that the equilibrium constant K is large:

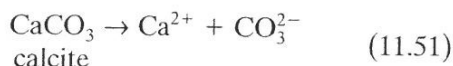
$$\frac{[A_2 B_3]}{[A]^2 [B]^3} = K \quad (11.49)$$

If, however, ΔG_R° for reaction 11.17 in the standard state is positive, the reaction runs the other way:



Therefore, the equilibrium constant is small because the reactants are more abundant than the product. Evidently, the sign and magnitude of ΔG_R° of a reaction can be used to predict both the *direction* in which a reaction runs in the standard state and the magnitude of the equilibrium constant.

In conclusion, we calculate ΔG_R° for the familiar reaction:



using the data in Appendix B:

$$\begin{aligned} \Delta G_R^\circ &= [(-132.3) + (-126.17)] - [(-269.9)] \\ &= +11.43 \text{ kcal} \end{aligned} \quad (11.52)$$

Since ΔG_R° is positive, the reaction runs from right to left in the standard state and calcite precipitates. The equilibrium constant for reaction 11.51 should therefore be a small number, as we know it to be. It is not surprising that the reaction (equation 11.51) runs to the left in the standard state because the activities of the ions in the standard state are each equal to one and the ion activity product is therefore much larger than the equilibrium constant.

We can also calculate ΔH_R° and ΔS_R° for reaction 11.51 using values of H_f° from Appendix B and of S_f° from Krauskopf (1979):

$$\begin{aligned} \Delta H_R^\circ &= [(-129.7) + (-161.84)] - [(-288.6)] \\ &= -2.94 \text{ kcal} \end{aligned} \quad (11.53)$$

$$\begin{aligned} \Delta S_R^\circ &= [(-12.7) + (-13.6)] - [(+22.2)] \\ &= -48.5 \text{ cal/deg} \end{aligned} \quad (11.54)$$

Note that ΔS_R° for the dissociation of calcite into ions is negative, implying a *decrease* in the entropy. Note also that the entropy change is expressed in calories per degree (cal/deg) rather than in kilocalories (kcal). We can now combine ΔH_R° and ΔS_R° for reaction 11.51 and calculate ΔG_R° , thereby testing the internal consistency of the thermodynamic data:

$$\Delta G_R^\circ = -2.94 - \frac{298.15(-48.5)}{1000} \quad (11.55)$$

where 298.15 is the temperature of the standard state and we divide the $T\Delta S$ term by 1000 in order to convert calories to kilocalories.

The result is:

$$\Delta G_R^\circ = +11.5 \text{ kcal} \quad (11.56)$$

which is in satisfactory agreement with $\Delta G_R^\circ = +11.43$ kcal calculated from equation 11.52. Note that ΔG_R° for the dissociation of calcite into ions is positive even though the reaction is exothermic, that is, ΔH_R° is negative. This occurs because more heat energy is consumed to decrease the entropy than is produced by the dissociation. Therefore, in this case ΔH_R° is *not* a reliable predictor of the direction in which the reaction runs in the standard state.

The sign of ΔH_R° does, however, enable us to use Le Châtelier's principle to predict how the equilibrium constant of reaction 11.51 varies with temperature. Since the reaction is exothermic as written, it produces heat which increases the temperature. Therefore, if reaction 11.51 is at equilibrium and the temperature is *decreased*, the reaction moves to the right because that tends to *increase*

the temperature. When equilibrium is reestablished at a lower temperature, the equilibrium constant is *larger* than it was before. Therefore, the solubility-product constant of calcite *increases* with *decreasing* temperature and calcite becomes *more* soluble. As we discussed in Section 10.4, thermodynamics provides a rational explanation for the temperature dependence of the solubility of calcite and allows us to predict the course of other chemical reactions provided the necessary data are available.

11.8 Derivation of the Law of Mass Action

Because the Gibbs free energy is defined in terms of both the first and second laws of thermodynamics, we can express it in terms of directly measurable properties. We can replace H in equation 11.46 by equation 11.13 and obtain:

$$G = E + PV - TS \quad (11.57)$$

For an infinitesimal change in state at constant temperature $dT = 0$ and therefore:

$$dG = dE + PdV + VdP - TdS \quad (11.58)$$

Substituting equation 11.8 for dE yields:

$$dG = dq - PdV + PdV + VdP - TdS \quad (11.59)$$

Since $dS = dq/T$, $dq = TdS$ and equation 11.59 therefore reduces to:

$$dG = VdP \quad (11.60)$$

We can use equation 11.60 to calculate the change in the Gibbs free energy of one mole of an ideal gas as a function of pressure at constant temperature. Since the volume of gases varies with pressure, we must express the volume in equation 11.60 by an equation that relates volume to pressure. According to the ideal gas law, for one mole of gas:

$$V = \frac{RT}{P} \quad (11.61)$$

where R is the gas constant.

Substituting into equation 11.60 yields:

$$dG = RT \left(\frac{dP}{P} \right) \quad (11.62)$$

Integrating from the standard pressure (P°) to some other pressure (P):

$$\int_{P^\circ}^P dG = RT \int_{P^\circ}^P \frac{dP}{P} \quad (11.63)$$

yields:

$$G_P - G^\circ = RT(\ln P - \ln P^\circ) \quad (11.64)$$

Since $P^\circ = 1 \text{ atm}$, $\ln P^\circ = 0$. Therefore, equation 11.64 reduces to:

$$G_P - G^\circ = RT \ln P \quad (11.65)$$

where G_P is the molar free energy of the ideal gas at pressure P , G° is the molar free energy in the standard state, and T is the temperature of the standard state in kelvins. The molar free energy is also called the chemical potential (μ), and equation 11.65 can be written as:

$$\mu = \mu^\circ + RT \ln P \quad (11.66)$$

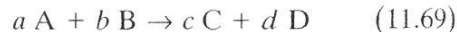
The free energy of n moles of an ideal gas can be calculated by multiplying equation 11.65 by n :

$$nG_P - nG^\circ = nRT \ln P \quad (11.67)$$

According to Dalton's law, the total pressure of a mixture of ideal gases is the sum of the partial pressures exerted by each gas in the mixture. Therefore, for a mixture of i ideal gases:

$$\sum n_i G_i - \sum n_i G_i^\circ = RT \sum n_i \ln P_i \quad (11.68)$$

Next, we assume that the mixture of ideal gases is the result of a chemical reaction and specify that the molar coefficients (n) of the *products* are *positive* and those of the *reactants* are *negative*. The provision is not new or arbitrary but goes back to the convention regarding the algebraic sign of heat. If we choose to represent the chemical reaction by the equation:



then we can evaluate what equation 11.68 tells us to do. For example, $\sum n_i G_i$ means:

$$\sum n_i G_i(\text{products}) - \sum n_i G_i(\text{reactants}) = \Delta G_R \quad (11.70)$$

Similarly, $\sum n_i G_i^\circ$ means:

$$\sum n_i G_i^\circ(\text{products}) - \sum n_i G_i^\circ(\text{reactants}) = \Delta G_R^\circ \quad (11.71)$$

and $RT \sum n_i \ln P_i$ means:

$$\begin{aligned} RT(c \ln P_C + d \ln P_D - a \ln P_A - b \ln P_B) \\ = RT(\ln P_C^c P_D^d - \ln P_A^a P_B^b) \\ = RT \ln \left(\frac{P_C^c P_D^d}{P_A^a P_B^b} \right) = RT \ln Q \end{aligned} \quad (11.72)$$

where Q is the *reaction quotient*. Therefore, we now restate equation 11.68 as:

$$\Delta G_R - \Delta G_R^\circ = RT \ln Q \quad (11.73)$$

which is valid for a reacting mixture of ideal gases.

The reaction quotient (Q) varies continuously as the reaction proceeds until the reaction reaches a state of equilibrium. After equilibrium has been established, the partial pressures of the reactant and product gases become constant because the forward and backward reactions are occurring at the same rate. Therefore, at equilibrium the reaction quotient becomes constant and is known to us as the *equilibrium constant*.

Next, we consider a very *important point*: the free energy change of a chemical reaction *at equilibrium* is equal to zero. This statement becomes obvious if we recall that ΔG_R is regarded as the driving force of a chemical reaction. When a reaction is in equilibrium, there is no driving force pushing it one way or the other. However, note that ΔG_R° is *not* zero because it is the change in the Gibbs free energy when the reaction takes place in the standard state. Therefore, ΔG_R° is a *constant*, whereas ΔG_R varies as the reaction proceeds from the standard state toward equilibrium.

When the reaction reaches equilibrium, $\Delta G_R \rightarrow 0$, but ΔG_R° remains the same.

Returning to equation 11.73 we can now state that for a mixture of ideal gases at equilibrium $\Delta G_R = 0$ and therefore:

$$\Delta G_R^\circ = -RT \ln K \quad (11.74)$$

where K is the equilibrium constant at the standard temperature. This remarkable equation establishes the functional relationship that exists between ΔG_R° and K , both of which are constants that apply to chemical reactions in the standard state and at equilibrium, respectively. In the process of deriving this relationship we have also proved the validity of the Law of Mass Action for reactions among ideal gases. These two results are among the most important contributions thermodynamics has made to science. Equation 11.74 allows us to calculate the numerical values of equilibrium constants at 25°C for all chemical reactions for which we can calculate ΔG_R° from existing measurements of the standard free energies of formation.

We reduce equation 11.74 to a formula by converting the equation to logarithms to the base 10 and by substituting $R = 1.987$ cal/deg · mol and $T = 298.15$ K:

$$\Delta G_R^\circ = -\frac{2.3025 \times 1.987 \times 298.15}{1000} \log K \quad (11.75)$$

$$= -1.364 \log K \quad (11.76)$$

The right side of equation 11.75 must be divided by 1000 because R is expressed in calories, whereas ΔG_R° is in kilocalories. Therefore, the equilibrium constant K is related to ΔG_R° by the formula:

$$K = 10^{-\Delta G_R^\circ / 1.364} \quad (11.77)$$

This is a very useful result, but it applies only to reactions among ideal gases unless we take steps to make it applicable to real gases and to ions and molecules in aqueous solutions.

11.9 Fugacity and Activity

In order to make equations 11.74 and 11.77 applicable to reactions among *real gases* we define the *fugacity* as the partial pressure a real

gas would have if it were ideal. Therefore, fugacities satisfy equation 11.74. In order to apply equation 11.74 to *real solutions* of any kind we define the *activity* (a) of a solute as:

$$a = \frac{f}{f^\circ} \quad (11.78)$$

where f is the fugacity of the solute in the solution and f° is its fugacity when the vapor is in equilibrium with the pure substance in the standard state. Since fugacities are measured in atmospheres or other units of pressure, the activity as defined by equation 11.78 is a dimensionless number. It also follows from equation 11.78 that for a pure substance:

$$a = 1.0 \quad (11.79)$$

because the fugacity f of a pure substance is equal to f° .

Binary mixtures of two nonelectrolytes A and B obey Raoult's law, which states that the fugacity of component A is given by:

$$f_A = f_A^\circ N_A \quad (11.80)$$

where N_A is the mole fraction of A defined as:

$$N_A = \frac{n(A)}{n(A) + n(B)} \quad (11.81)$$

where $n(A)$ and $n(B)$ are the numbers of moles of A and B, respectively. When the solution is dilute, that is, $n(A) \ll n(B)$, the vapor obeys Henry's law:

$$f_A = k N_A \quad (11.82)$$

By dividing both sides of equation 11.82 by f_A° , we obtain:

$$\frac{f_A}{f_A^\circ} = \frac{k}{f_A^\circ} N_A = a_A \quad (11.83)$$

Next, we note that for a dilute solution of solute A in solvent B the number of moles of A is much smaller than the number of moles of B. Therefore, if $n(A) \ll n(B)$, equation 11.81 can be restated to a very good approximation as:

$$N_A = \frac{n(A)}{n(B)} \quad (11.84)$$

The mole fraction of A in a dilute solution can be transformed into the molal concentration of A (m_A) by converting $n(B)$ to kilograms of B:

$$m_A = \frac{n(A) \times 1000}{n(B) \times MW(B)} \quad (11.85)$$

where $MW(B)$ is the gram-molecular weight of B. Since $n(A)/n(B) = N_A$ in dilute solutions of A in B:

$$m_A = \frac{N_A \times 1000}{MW(B)} = N_A k' \quad (11.86)$$

and k' is a constant equal to $1000/MW(B)$. When we substitute equation 11.86 into equation 11.83, we obtain:

$$a_A = \frac{k}{f_A^\circ} N_A = \frac{k m_A}{f_A^\circ k'} \quad (11.87)$$

By combining the three constants in equation 11.87 we obtain:

$$a_A = \gamma_A m_A \quad (11.88)$$

where γ_A is the *activity coefficient*.

This relationship applies to all solutions regardless of the vapor pressures of the solutes or of the solvent. If the solution is a mixture of volatile nonelectrolytes such as acetone dissolved in water, we have no difficulty detecting the presence of acetone vapor and the definition of the activity of acetone in the solution given by equation 11.78 makes sense. Ions produced by the dissociation of salts, acids, and bases in water also produce vapors, but their partial pressures or fugacities are so low that they may be undetectable. However, this experimental difficulty does not alter the validity of relationships derived by consideration of vapors in equilibrium with solutes and solvents regardless of whether they are electrolytes or nonelectrolytes.

The concept of *activity* presented above is quite different from that discussed in Sections 9.2 and 11.3. We can reconcile the two representations by saying that activities are introduced in thermodynamics in order to extend the Law of Mass Action from reacting mixtures of ideal gases to chemical reactions among molecules and ions in nonideal aqueous solutions. This is accomplished

by means of activity coefficients that convert molal concentrations of ions and molecules into activities. When activity coefficients are regarded as dimensionless conversion factors, then activities have the same units as concentrations.

11.10 The van't Hoff Equation

Thermodynamics enables us to calculate the equilibrium constants of chemical reactions at 25°C. In order to extrapolate the equilibrium constant from the standard temperature to other temperatures between 0 and 100°C, we return briefly to equation 11.74:

$$\Delta G_R^\circ = -RT \ln K$$

By solving this equation for $\ln K$ we obtain:

$$\ln K = -\frac{\Delta G_R^\circ}{RT} \quad (11.89)$$

Next, we differentiate equation 11.89 with respect to T :

$$\frac{d \ln K}{dT} = -\left(\frac{1}{R}\right) d\left(\frac{\Delta G_R^\circ}{dT}\right) \quad (11.90)$$

Since $\Delta G_R^\circ = \Delta H_R^\circ - T\Delta S_R^\circ$, it follows that:

$$\frac{\Delta G_R^\circ}{T} = \frac{\Delta H_R^\circ}{T} - \Delta S_R^\circ \quad (11.91)$$

and:

$$d\frac{\Delta G_R^\circ}{dT} = -\frac{\Delta H_R^\circ}{T^2} - 0 \quad (11.92)$$

Substituting into equation 11.90 yields:

$$\frac{d \ln K}{dT} = \left(-\frac{1}{R}\right)\left(-\frac{\Delta H_R^\circ}{T^2}\right) = \frac{\Delta H_R^\circ}{RT^2} \quad (11.93)$$

Multiplying by dT gives:

$$d \ln K = \left(\frac{\Delta H_R^\circ}{R}\right)\left(\frac{dT}{T^2}\right) \quad (11.94)$$

We integrate equation 11.94 from the standard temperature (T°) to some other temperature (T), which may be higher or lower than T° :

$$\int_{T^\circ}^T d \ln K = \frac{\Delta H_R^\circ}{R} \int_{T^\circ}^T \left(\frac{dT}{T^2}\right) \quad (11.95)$$

The result of the integration is:

$$\ln K_T - \ln K_{T^\circ} = \left(-\frac{\Delta H_R^\circ}{R}\right)\left(\frac{1}{T} - \frac{1}{T^\circ}\right) \quad (11.96)$$

Equation 11.96 is the important van't Hoff equation, which we now restate in terms of logarithms to the base 10:

$$\log K_T = \log K_{T^\circ} - \frac{\Delta H_R^\circ}{2.3025R}\left(\frac{1}{T} - \frac{1}{298.15}\right) \quad (11.97)$$

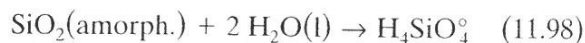
Note that equation 11.97 can be used to calculate K at any temperature between 0 and 100°C if we know the value of K at the standard temperature. However, if the reference temperature *differs* from 25°C, then ΔH_R° must be evaluated at that temperature by means of heat capacities, as shown in equation 11.32 or 11.35.

We have now completed this selective presentation of thermodynamics, which has enabled us to prove the validity of the Law of Mass Action and to calculate equilibrium constants of chemical reactions not only at the standard temperature but also at other temperatures between 0 and 100°C.

11.11 Solubility of Amorphous Silica between 0 and 100°C

At the end of Chapter 9 we calculated the solubility of amorphous silica in water as a function of pH at 25°C. We are now in a position to investigate how the solubility of amorphous silica varies as a function of temperature at different values of the pH. The thermodynamic constants required for this calculation are listed in Appendix B.

Amorphous silica reacts with water to form silicic acid:



The change in Gibbs free energy of this reaction in the standard state is:

$$\begin{aligned} \Delta G_R^\circ &= [(-312.66)] - [(-203.33) + 2(-56.687)] \\ &= +4.044 \text{ kcal} \end{aligned} \quad (11.99)$$

and the equilibrium constant K_T is:

$$K_{T^0} = 10^{-(4.044/1.364)} = 10^{-2.96} \quad (11.100)$$

The change in enthalpy in the standard state is:

$$\begin{aligned} \Delta H_R^\circ &= -349.1 - [-215.94 + 2(-68.315)] \\ &= +3.47 \text{ kcal} \end{aligned} \quad (11.101)$$

The reaction is *endothermic*, suggesting that its equilibrium constant *increases* with temperature. From the van't Hoff equation (11.97):

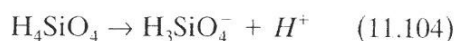
$$\begin{aligned} \log K_T &= \\ -2.96 - \frac{3.47 \times 10^3}{2.3025 \times 1.987} \left(\frac{1}{T} - \frac{1}{298.15} \right) \end{aligned} \quad (11.102)$$

where $R = 1.987 \text{ cal/deg} \cdot \text{mol}$. Note that ΔH_R° must be converted to calories in order to be compatible with R . Equation 11.102 reduces to:

$$\log K_T = -2.96 - 758.4 \left(\frac{1}{T} - 0.00335 \right) \quad (11.103)$$

When $T > 298.15 \text{ K}$, $1/T - 0.00335$ is negative and the algebraic sign of the second term in equation 11.103 changes to plus. Therefore, at $T = 273.15 \text{ K}$ (0°C) equation 11.103 yields $K_T = 10^{-3.20}$, whereas at 373.15 K (100°C) $K_T = 10^{-2.45}$. Evidently, the equilibrium constant of reaction 11.98 *increases* with *increasing* temperature as predicted.

The first dissociation of silicic acid forms H_3SiO_4^- and H^+ :



The standard free energy change is $\Delta G_R^\circ = +13.24 \text{ kcal}$, and the standard enthalpy change is $\Delta H_R^\circ = +7.0 \text{ kcal}$. Note that the G_f° and H_f° of H^+ are both equal to zero. The equilibrium constant of reaction 11.104 at 25°C is $K = 10^{-9.71}$ and varies with temperature in accordance with the van't Hoff equation:

$$\log K_T = -9.71 - 1530.03 \left(\frac{1}{T} - 0.00335 \right) \quad (11.105)$$

Consequently at 0°C , $K_T = 10^{-10.19}$, whereas at 100°C , $K_T = 10^{-8.68}$. Once again, the equilibrium constant *increases* with rising temperature because reaction 11.104 is endothermic.

These calculations illustrate the point that thermodynamics not only confirms our qualitative predictions based on Le Châtelier's principle but actually permits us to determine the numerical values of the equilibrium constants of all kinds of chemical reactions between temperatures of 0 and 100°C or other appropriate temperature limits.

Although this is a very important advance in geochemistry, the calculations can become quite tedious. Therefore, it is helpful to summarize the results of thermodynamic calculations by means of suitable diagrams. In the case at hand, we want to depict the concentration of SiO_2 in saturated solutions of amorphous silica at different values of the pH and temperature. We therefore carry out the calculations first made in Section 9.6 at a series of temperatures for selected pH values and plot the results in Figure 11.1 as a set of curves in coordinates of the concentration of SiO_2 and temperature. The diagram depicts the increase in the solubility of amorphous SiO_2 with increasing temperature at specified values of the pH. Such diagrammatic representations of chemical reactions are widely used in geochemistry and therefore are the subject of Chapter 12.

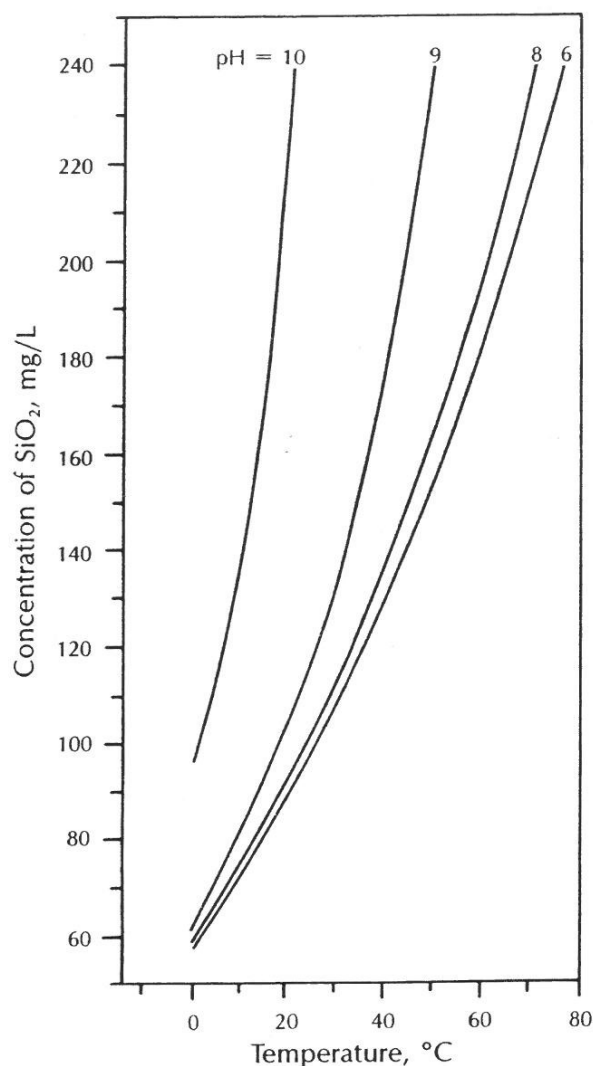
11.12 Summary

Thermodynamics is based on three statements regarding the relationship between heat and other forms of energy that do not require proof because they are self-evident. Starting from these basic premises, thermodynamics has grown into a large body of knowledge by rigorous deductive reasoning aided by mathematics. The products of this development include the concepts of enthalpy and Gibbs free energy and a proof of the Law of Mass Action for reacting mixtures of ideal gases.

These results of thermodynamics have great practical value in geochemistry because they permit us to calculate the equilibrium constants of

chemical reactions at any desired temperature, limited only by the availability of the necessary thermodynamic parameters and constrained by the stabilities of the reactants and products. Consequently, we can now explore the behavior of chemical reactions on the surface of the Earth and use the results to specify the stability limits of the common rock-forming minerals. Many of the reactions that take place in the natural environment can be studied in the laboratory only with great difficulty because these reactions are so slow. Thermodynamics enables us to make predictions about the outcome of these reactions that can be tested by evidence in the field.

Figure 11.1 Solubility of amorphous silica as a function of temperature at different pH values. The equilibrium constants at 25°C were calculated from equation 11.77 using the standard free energies of formation from Appendix B. Equilibrium constants at other temperatures were derived from the van't Hoff equation (11.97) based on K and ΔH_R° . The activity coefficients of all ions and molecules were assumed to be equal to one. The resulting curves indicate that the concentration of SiO_2 in milligrams per liter in saturated solutions of amorphous silica in water increases both with increasing temperature and pH.



Problems

1. Calculate the enthalpy change when fluorite (CaF_2) dissolves in water in the standard state.
2. Based on the result in Problem 1, predict how the solubility of fluorite varies with temperature.
3. Calculate the standard free energy change of the reaction by which fluorite dissociates into ions.
4. Calculate the solubility of fluorite in water at 10, 20, and 30°C and express each in terms of the concentration of Ca^{2+} in mg/L. (Assume $\gamma = 1$ for all species.)
5. Calculate K_T for the reaction between albite ($\text{NaAlSi}_3\text{O}_8$), water, and H^+ to form kaolinite ($\text{Al}_2\text{Si}_2\text{O}_5(\text{OH})_4$) and silicic acid.
6. Apply the Law of Mass Action to the reaction in Problem 5, take logarithms to the base 10, and express the resulting equation in terms of $\log [\text{Na}^+]/[\text{H}^+]$ and $\log [\text{H}_4\text{SiO}_4]$.
7. The water of the Mississippi River at Baton Rouge, Louisiana, has the following chemical composition at $\text{pH} = 7.2$.

	ppm		ppm
HCO ₃ ⁻	101	Mg ²⁺	7.6
SO ₄ ²⁻	41	Na ⁺	11
Cl ⁻	15	K ⁺	3.1
NO ₃ ⁻	1.9	SiO ₂	5.9
Ca ²⁺	34		

Calculate $\log [\text{Na}^+]/[\text{H}^+]$ and $\log [\text{H}_4\text{SiO}_4]$ for this water using the necessary activity coefficients.

8. Use the results of Problems 6 and 7 to determine the direction of the reaction in Problem 5 in the

Mississippi River at Baton Rouge; that is, determine whether albite or kaolinite is produced.

9. Make the necessary calculation (assuming $T = 25^\circ\text{C}$) to determine whether the Mississippi River at Baton Rouge is saturated or undersaturated with respect to (a) calcite, (b) gypsum, (c) amorphous silica, (d) quartz, and (e) dolomite.

10. Calculate the value of G_f° for BaF_2 given that the concentration of Ba^{2+} in a saturated solution at 25°C , having an ionic strength of 0.01, is 1.12×10^{-2} m/L. Assume that Ba^{2+} and F^- are derived entirely from BaF_2 and that hydrolysis is negligible.

References

- ANDERSON, G. M., 1996. *Thermodynamics of Natural Systems*. Wiley, New York, 382 pp.
- BARIN, I., F. SAUERT, E. SCHULTZE-RHONHOF, and W. S. SHENG, 1989. *Thermochemical Data of Pure Substances*, VCH Publishers, New York, 1739 pp.
- FRASER, D. G. (Ed.), 1977. *Thermodynamics in Geology*. Reidel, Amsterdam, 410 pp.
- KELLEY, K. K., 1962. Heats of free energies of formation of anhydrous silicates. *U.S. Bureau of Mines Rept. Invest.* 5901.
- KERN, R., and A. WEISBROD, 1967. *Thermodynamics for Geologists*. W. H. Freeman, New York, 304 pp.
- KRAUSKOPF, K. B., 1979. *Introduction to Geochemistry*, 2nd ed. McGraw-Hill, New York, 617 pp.
- MOORE, W. J., 1955. *Physical Chemistry*, 2nd ed. Prentice-Hall, Upper Saddle River, NJ, 633 pp.
- NORDSTROM, D. K., and J. L. MUNOZ, 1986. *Geochemical Thermodynamics*. Blackwell Scientific, Palo Alto, 477 pp.
- RIFKIN, J., 1980. *Entropy*. Bantam, New York, 302 pp.
- ROBIE, R. A., B. S. HEMINGWAY, and J. R. FISHER, 1978. Thermodynamic properties of minerals and related substances at 298.15 K and 1 bar pressure and at higher temperatures. *U.S. Geol. Survey Bull.* 1452.
- WEAST, R. C., M. J. ASTLE, and W. H. BEYER (Eds.), 1986. *CRC Handbook of Chemistry and Physics*. CRC Press, Boca Raton, FL.
- WOOD, B. J., and D. G. FRASER, 1976. *Elementary Thermodynamics for Geologists*. Oxford University Press, Oxford, England, 303 pp.
- WOODS, T. L., and R. S. GARRELS, 1987. *Thermodynamic Values at Low Temperature for Natural Inorganic Materials*. Oxford University Press, New York, 242 pp.

Mineral Stability Diagrams

When minerals participate in chemical reactions as reactants or products, they are either consumed or produced depending on the direction in which the reaction is moving. We can therefore say that *reactant* minerals are *unstable*, whereas *product* minerals are *stable* while the reaction is in progress. If the reaction ultimately reaches equilibrium, *both* reactant and product minerals become stable. Thus, we can investigate the stability of minerals under two kinds of conditions:

1. We can determine the environmental conditions necessary for minerals to coexist at equilibrium.
2. We can determine which minerals are stable and which are unstable in a particular geochemical environment.

The theoretical basis for these kinds of studies is provided by the Law of Mass Action and by the relationship between the standard free energy change of a chemical reaction and its equilibrium constant at 25°C.

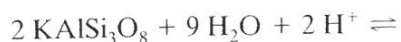
The activity diagrams we want to construct are based on the *incongruent* solution of certain aluminosilicate minerals in igneous and high-grade metamorphic rocks. The reactions on which the diagrams are based contribute to chemical weathering and result in the formation of oxides, hydroxides, clay minerals, or zeolites, depending on the geochemical environment. Transformations among minerals that form only at elevated temperature and pressure are inappropriate in activity diagrams drawn at the tem-

perature and pressure of the Earth's surface. However, the stability of minerals can also be defined by equations arising from their *congruent* solution in dilute aqueous solutions. Both kinds of activity diagrams present complex chemical interactions in a form that is especially appropriate for geologists accustomed to using phase diagrams in igneous petrology.

12.1 Chemical Weathering of Feldspars

The feldspars in igneous rocks crystallize from silicate melts at high temperature, but can also form on the surface of the Earth by reactions of aqueous solutions with other aluminosilicates such as the clay minerals. In Section 10.6 we discussed how to construct the equations that represent the reactions among aluminosilicates, but we were unable to proceed because we did not know the equilibrium constants of the reactions we wanted to investigate.

Now we can restate the equation for the conversion of microcline to kaolinite:



Using the standard free energies of formation from Appendix B we obtain:

$$\begin{aligned} \Delta G_R^\circ &= [-(906.84) + 2(-67.70) + 4(-312.66)] \\ &\quad - [2(-894.9) + 9(-56.687)] \\ &= +7.103 \text{ kcal} \end{aligned} \quad (12.2)$$

Therefore, the equilibrium constant of the reaction at 25°C is:

$$K = 10^{-(7.103/1.364)} = 10^{-5.21} \quad (12.3)$$

Applying the Law of Mass Action when the reaction is at equilibrium gives:

$$\frac{[\text{H}_4\text{SiO}_4]^4[\text{K}^+]^2}{[\text{H}^+]^2} = 10^{-5.21} \quad (12.4)$$

where the square brackets signify activities or fugacities of reactants and products. Next, we take logarithms of equation 12.4:

$$4 \log [\text{H}_4\text{SiO}_4] + 2 \log [\text{K}^+] - 2 \log [\text{H}^+] = -5.21 \quad (12.5)$$

Collecting terms:

$$4 \log [\text{H}_4\text{SiO}_4] + 2(\log [\text{K}^+] - \log [\text{H}^+]) = -5.21 \quad (12.6)$$

from which it follows that:

$$4 \log [\text{H}_4\text{SiO}_4] + 2 \log \frac{[\text{K}^+]}{[\text{H}^+]} = -5.21 \quad (12.7)$$

and:

$$\log \frac{[\text{K}^+]}{[\text{H}^+]} = -2 \log [\text{H}_4\text{SiO}_4] - 2.60 \quad (12.8)$$

Equation 12.8 is a straight line, shown in Figure 12.1, along which $\log [\text{K}^+]/[\text{H}^+]$ and $\log [\text{H}_4\text{SiO}_4]$ have the values required for equilibrium between microcline and kaolinite. If the two parameters have values that place the environment *off* the line, the reaction is *not* at equilibrium. In that case, we can determine the direction in which the reaction proceeds by applying Le Châtelier's principle to equation 12.1. If $\log [\text{H}_4\text{SiO}_4]$ of the environment is *greater* than required for equilibrium at a particular value of $\log [\text{K}^+]/[\text{H}^+]$, reaction 12.1 moves to the left, thereby converting kaolinite into microcline. Therefore, microcline is *stable* under these conditions, whereas kaolinite is *unstable*. The same reasoning indicates that if $\log [\text{H}_4\text{SiO}_4]$ is *less* than required for equilibrium, kaolinite is stable and microcline is unstable at 25°C.

We might have chosen to write an equation for the conversion of microcline to gibbsite instead of kaolinite. The resulting line plots to the left of the microcline-kaolinite equilibrium in Figure 12.1. This means that microcline is reacting to form gibbsite under conditions in which kaolinite, not microcline, is stable. Consequently, the transformation of microcline to gibbsite is *metastable* and must be excluded from the diagram.

When reaction 12.1 occurs in a closed environment having a small water/rock ratio, $\log [\text{H}_4\text{SiO}_4]$ and $\log [\text{K}^+]/[\text{H}^+]$ may change as the reaction progresses until equilibrium is established. In that case, kaolinite and microcline coexist in equilibrium with the aqueous solution, provided the temperature remains constant and the activities of the ions and molecules remain unchanged. On the other hand, if the environment is open or the water/rock ratio is large, the reaction may not achieve equilibrium because it is unable to change the activity of H_4SiO_4 and the $[\text{K}^+]/[\text{H}^+]$ ratio of the environment enough to make them compatible with equation 12.8. In this case, the reaction continues until either kaolinite or microcline is used up, depending on the direction in which it is moving; that is, the reaction goes to *completion*.

We can investigate the effect of a temperature change on reaction 12.1 by calculating ΔH_R° :

$$\begin{aligned} \Delta H_R^\circ &= [(-983.5) + 2(-60.32) + 4(-348.3)] \\ &\quad - [2(-948.7) + 9(68.315)] \\ &= +14.895 \text{ kcal} \end{aligned} \quad (12.9)$$

We see that the reaction is *endothermic* when it occurs in the standard state, which means that it consumes heat in the forward direction. Therefore, an *increase* in temperature favors the conversion of microcline to kaolinite and causes the equilibrium constant to *increase*. The relationship between K_{T° and K_T is given by the van't Hoff equation:

$$\log K_T = \log K_{T^\circ} - \frac{\Delta H_R^\circ \times 10^3}{2.3025R} \left(\frac{1}{T} - 0.00335 \right) \quad (12.10)$$

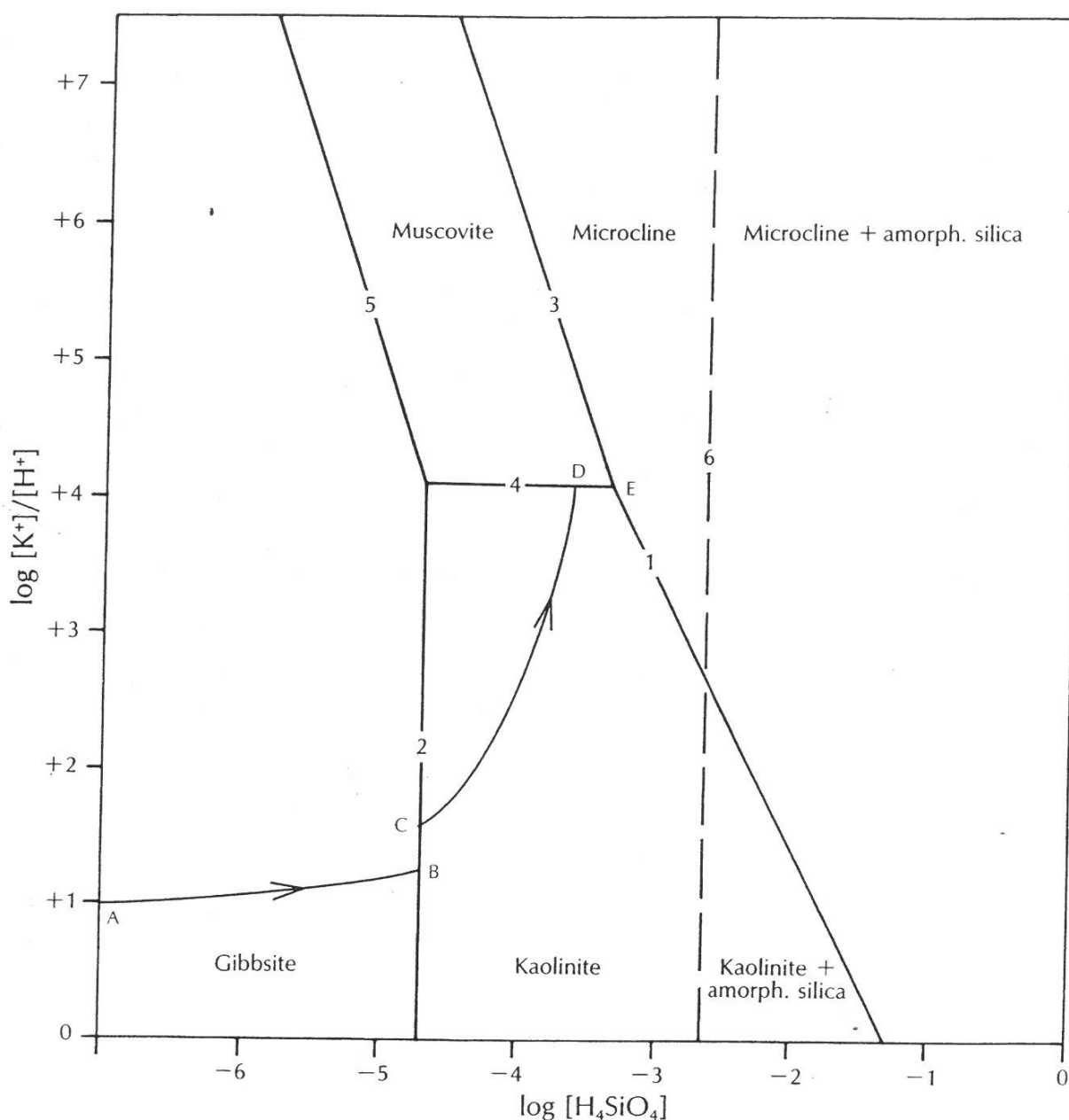


Figure 12.1 Stabilities of microcline, muscovite (proxy for illite), kaolinite, gibbsite, and amorphous silica in equilibrium with K^+ , H^+ , and H_4SiO_4 in aqueous solution at 25°C and 1 atm pressure. The equations for the stability boundaries are based on the standard free energy values from Appendix B and are listed in Table 12.1. Note that K-smectite, pyrophyllite, and quartz are omitted from the diagram in order to keep it simple. Note also that the diagram lies in the upper left quadrant of the coordinate system. The reaction path was adapted from Steinmann et al. (1994).

Since $K_{T^0} = 10^{-5.21}$, $\Delta H_R^\circ = +14.895$ kcal, and if $T = 308.15$ K (35°C):

$$\log K_T = -5.21 - \frac{14,895}{4,575} (0.00324 - 0.00335)$$

$$K(35^\circ\text{C}) = 10^{-4.85} \quad (12.11)$$

Since $10^{-4.85} > 10^{-5.21}$, the equilibrium *does* shift to the right in favor of kaolinite with increasing temperature, as predicted.

When reaction 12.1 proceeds from left to right as written, it produces K^+ ions and H_4SiO_4 molecules while consuming H^+ and H_2O . Therefore, if the system is closed, the chemical composition of the water changes in such a way that $\log [\text{H}_4\text{SiO}_4]$ and $\log [\text{K}^+]/[\text{H}^+]$ become more compatible with the values required for equilibrium. Therefore, chemical weathering alters the chemical composition of water and hence its *quality* as a natural resource. Moreover, reactions such as 12.1 take place not only on the surface of the Earth but also *below* the surface where groundwater is in contact with minerals that are unstable in the presence of water. Chemical “weathering” is *not* restricted to places that are exposed to the weather, but occurs at varying depth below the surface depending on the location of the water table and the porosity and permeability of the rocks.

We can now complete the construction of the stability diagram for microcline (Figure 12.1) by adding other crystalline or amorphous compounds that can occur in the system composed of SiO_2 – Al_2O_3 – K_2O – HCl – H_2O . These include gibbsite, muscovite (proxy for illite), K-smectite, pyrophyllite, and amorphous silica. Note that “HCl” serves as a source of hydrogen ions and is not actually present in natural environments. Note also that the stability fields of K-smectite and pyrophyllite are not shown in Figure 12.1, but they do occur between kaolinite and microcline. The equations required to define the stability limits of the solid phases in this system are derivable by the same procedure we demonstrated above for reaction 12.1 and are therefore listed in Table 12.1.

The boundaries in Figure 12.1 are the loci of points whose coordinates have the values required

for equilibrium between the solids in adjoining fields. Each field is labeled to indicate which solid is stable within its boundaries. The solids are stable only within their stability fields. When they are exposed to an environment that lies outside their own stability field, they are converted into the solid that is stable in that environment. For example, if a grain of microcline is placed into an environment within the stability field of kaolinite, reaction 12.1 will run from left to right and convert the microcline into kaolinite, provided enough time is available to allow the reaction to run to completion. Similarly, if kaolinite is placed into an environment within the stability field of microcline, it forms secondary K-feldspar known as adularia (Kastner, 1971; Kastner and Siever, 1979; Mensing and Faure, 1983).

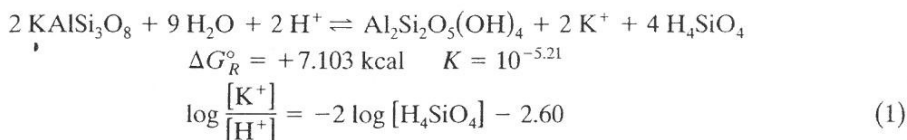
Most natural waters plot within the stability field of kaolinite in Figure 12.1, whereas seawater straddles the boundary between muscovite (illite) and microcline. Therefore, microcline and other feldspars in igneous and metamorphic rocks on the continents weather to form clay minerals (kaolinite and smectite). When these clay minerals are deposited in the oceans, they tend to take up K^+ and may be converted into illite (represented by muscovite in Figure 12.1) or low-temperature feldspar, given sufficient time.

The evolution of the chemical composition of water as a result of chemical weathering of minerals such as microcline can be predicted by numerical modeling (Helgeson et al., 1969). The progress of such reactions produces curved reaction paths on activity diagrams. An example of such a path is shown in Figure 12.1 based on modeling by Steinmann et al. (1994). Their calculations were based on the assumptions that:

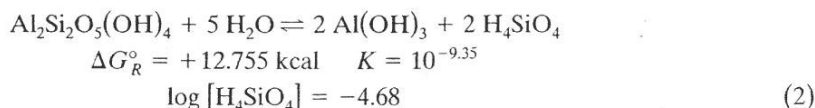
1. The system is closed and the water/rock ratio is small.
2. The secondary minerals are in equilibrium with ions in the solution even though the primary minerals are not in equilibrium.
3. Temperature and pressure remain constant.
4. Aluminum is conserved in the reactions.
5. The solutions remain dilute such that all activity coefficients are close to unity.

Table 12.1 Stability Relations Among Microcline, Kaolinite, Gibbsite, and Amorphous Silica in the Presence of Water at 25 °C and 1 atm pressure^a

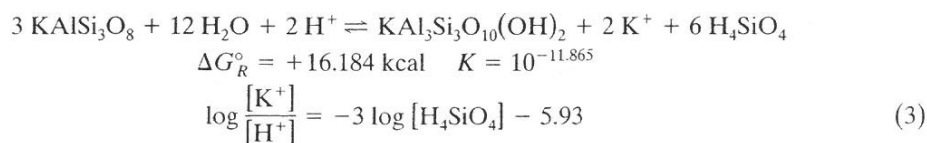
Microcline-kaolinite



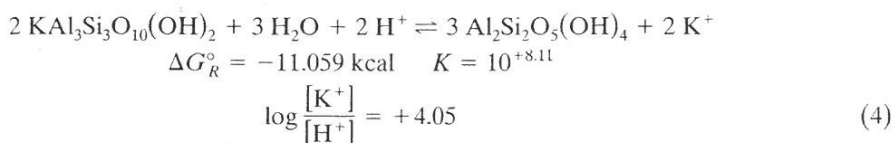
Kaolinite-gibbsite



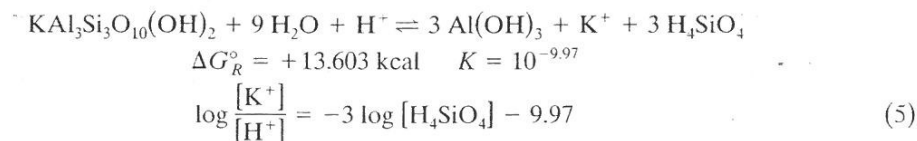
Microcline-muscovite



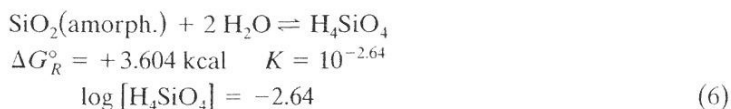
Muscovite-kaolinite



Muscovite-gibbsite

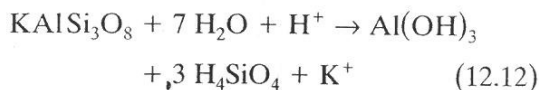


Solubility limit of amorphous silica



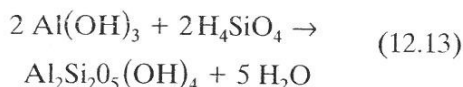
^aBased on the thermodynamic data in Appendix B.

When microcline is exposed to an environment represented by the coordinates of point *A* in Figure 12.1 (i.e., $\log [\text{H}_4\text{SiO}_4] = -6.0$, $\log [\text{K}^+]/[\text{H}^+] = -2.0$, $\text{pH} = 4$), it will react to form gibbsite:

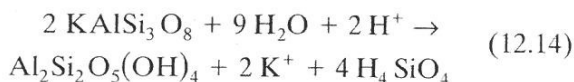


As a result, the $[\text{K}^+]/[\text{H}^+]$ ratio and $[\text{H}_4\text{SiO}_4]$ of the water both increase because the system is assumed to be closed and the water/rock ratio is small. Therefore, the chemical composition of the aqueous phase changes along the reaction path from *A* to *B*.

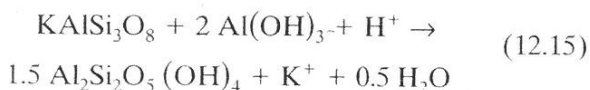
When the chemical composition of the aqueous phase reaches point *B*, kaolinite becomes the stable phase. Therefore, the previously formed gibbsite reacts with silicic acid in the water to form kaolinite:



This reaction consumes silicic acid, but does not affect the $[\text{K}^+]/[\text{H}^+]$ ratio of the water. At the same time, microcline continues to dissolve incongruently to form kaolinite (equation 1, Table 12.1):



This reaction releases silicic acid and also causes the $[\text{K}^+]/[\text{H}^+]$ ratio of the water to increase. If equilibrium between gibbsite and kaolinite is maintained, the concentration of silicic acid in the water must remain constant until all of the previously formed gibbsite has been converted to kaolinite. Under these conditions, the reactions at point *B* in Figure 12.1 can be represented by one equation written in such a way that Si is conserved (Steinmann et al., 1994):

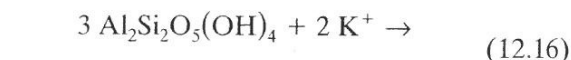


This equation can be constructed by multiplying equation 12.14 by 1/2 and adding it to 12.13. The

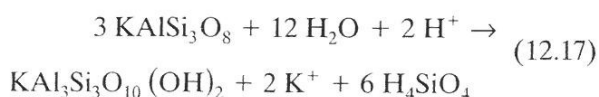
reactions represented by equation 12.15 cause the $[\text{K}^+]/[\text{H}^+]$ ratio of the water to rise along the gibbsite-kaolinite boundary from point *B* to point *C*, which is reached when all of the gibbsite has been converted to kaolinite.

Subsequently, microcline reacts with ions in the water to form kaolinite in accordance with equation 12.14. As a result, the $[\text{K}^+]/[\text{H}^+]$ ratio as well as $[\text{H}_4\text{SiO}_4]$ both increase along the reaction path between points *C* and *D*, provided that enough microcline is available to allow reaction 12.14 to continue.

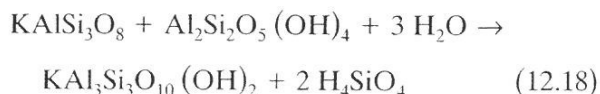
If the chemical composition of the water reaches the stability boundary between kaolinite and muscovite (illite), kaolinite becomes unstable and reacts with the ions in the water to form muscovite (equation 4, Table 12.1):



This reaction causes a decrease of the $[\text{K}^+]/[\text{H}^+]$ ratio of the water without affecting the activity of $[\text{H}_4\text{SiO}_4]$. However, if microcline is still available, it will now form muscovite (equation 3, Table 12.1):



This reaction increases both the $[\text{K}^+]/[\text{H}^+]$ ratio and $[\text{H}_4\text{SiO}_4]$. Since the composition of the water must remain on the stability boundary between kaolinite and muscovite until all of the previously formed kaolinite has been converted to muscovite, the $[\text{K}^+]/[\text{H}^+]$ ratio of the solution must remain constant. Therefore, equations 12.16 and 12.17 can be combined to yield:



This reaction increases $[\text{H}_4\text{SiO}_4]$ without affecting the $[\text{K}^+]/[\text{H}^+]$ ratio of the water. Therefore, the chemical composition of the water moves along the kaolinite-muscovite boundary in Figure 12.1 from point *D* toward point *E*.

If sufficient microcline is available to drive reaction 12.18 until the chemical composition of the water reaches point *E*, all three minerals (microcline, kaolinite, and muscovite) coexist in equilibrium with the ions in the solution. Consequently, the chemical weathering reactions stop at point *E* provided that the temperature remains constant and the system remains closed.

The reaction path shown in Figure 12.1 is one of an infinite number of possible paths whose course depends on the initial chemical composition and pH of the water. The work of Steinmann et al. (1994) indicated that the shape of reaction paths in Figure 12.1 is especially sensitive to the initial pH of the water. If $\text{pH} > 5$, the reaction paths proceed from point *A* to the gibbsite–muscovite boundary without entering the kaolinite field.

The purpose of this discourse is to emphasize that the reactions between microcline (or another mineral) and acidified water not only cause transformations among the solid phases but also affect the chemical composition of the water. These reactions therefore contribute significantly to the chemical evolution of water both on the surface and below the surface of the Earth. Therefore, activity diagrams like Figure 12.1 provide useful information about several geochemical questions:

1. What environmental conditions are required to allow a particular mineral to form?
2. What minerals are stable in a given geochemical environment?
3. What ions or molecules are consumed or produced when an unstable mineral reacts in a given geochemical environment?
4. How does the water evolve chemically when reactions occur in a closed system with a small water/rock ratio?

The answers to these kinds of questions provide information that is useful in hydrogeology, sedimentary petrology, soil science, geochemical prospecting, economic geology, clay mineralogy, and other important applications of inorganic aqueous geochemistry. A large number of activity

diagrams for many different systems at different temperatures and pressures was published by Bowers et al. (1984).

The stability limits of albite ($\text{NaAlSi}_3\text{O}_8$) and anorthite ($\text{CaAl}_2\text{Si}_2\text{O}_8$) are shown in Figure 12.2, and the relevant equations are listed in Table 12.2. These diagrams can be interpreted in the same way as the microcline diagram. They are, in fact, complementary to Figure 12.1 because they not only introduce Na^+ and Ca^{2+} , respectively, but also involve H^+ and H_4SiO_4 . In most natural environments all three minerals are likely to be present. Therefore, the activities of H^+ and H_4SiO_4 of natural solutions reflect the progress of all three chemical reactions occurring simultaneously. In addition, other reactions may be taking place involving other silicate minerals, as well as carbonates, phosphates, sulfides, or sulfates. As a result, predictions about the progress of one particular reaction occurring in such complex natural environments may be affected by the other reactions that are occurring simultaneously.

The stability diagram of albite in Figure 12.2A includes the Na-mica *paragonite* and therefore is similar to that of microcline. The solubility limit of magadiite (Surdam and Eugster, 1976) is listed in Table 12.2 but is not shown in Figure 12.2A. The stability of analcime is discussed in Section 12.2 with other zeolite minerals, whereas reactions among the clay minerals are discussed in Chapter 13.

The stability field of the Ca-mica *margarite* is not shown in Figure 12.2B because it is a rare mineral and its standard free energy of formation is not well known. In addition, the activity of Ca^{2+} in aqueous solutions reacting with anorthite is limited by the solubility of calcite, which depends on the fugacity of CO_2 of the environment and on the H^+ activity. The equation that governs the solubility limit of calcite can be constructed as follows. We know that calcite dissolves to form Ca^{2+} and HCO_3^- at pH values between 6.4 and 10.3 (Section 9.5). However, the bicarbonate ion is also in equilibrium with molecular H_2CO_3^0 , which equilibrates with $\text{CO}_2(\text{g}) + \text{H}_2\text{O}$. Therefore, we express the solubility of calcite in terms of Ca^{2+} and $\text{CO}_2(\text{g})$ because we

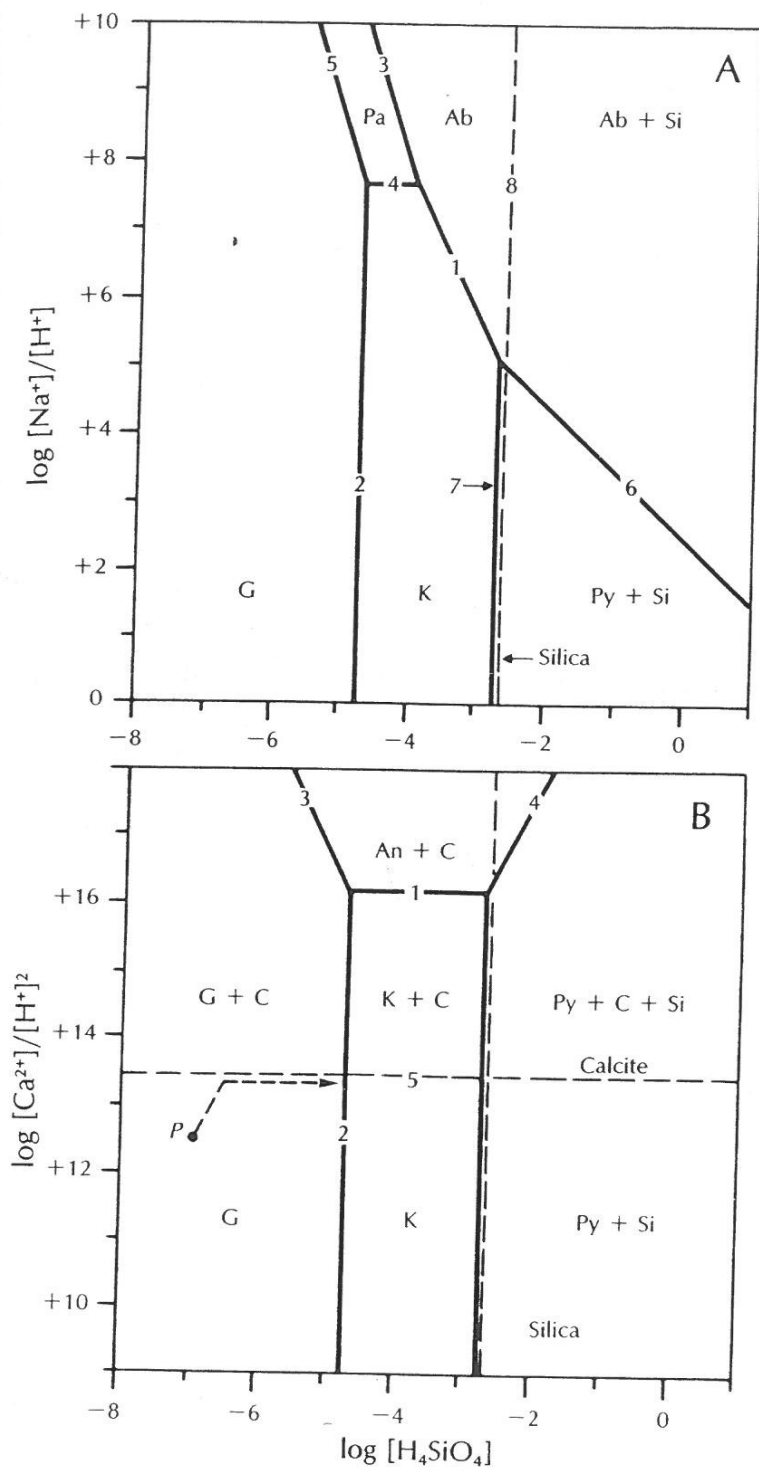
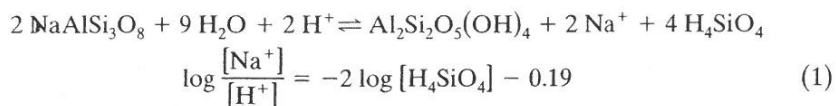


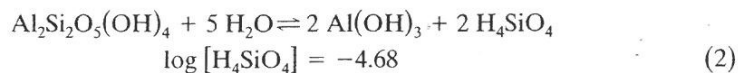
Figure 12.2 Stability of albite (**A**) and anorthite (**B**) in the presence of acidified water at 25°C and 1 atm pressure based on equations listed in Table 12.2. Ab = albite, K = kaolinite, G = gibbsite, Pa = paragonite, Py = pyrophyllite, Si = amorphous silica, An = anorthite, C = calcite. Na- and Ca-smectite and the solubility limits of quartz and magadiite are not shown, but the magadiite solubility limit is given by equation 9 in Table 12.2A. The solubility limit of calcite (equation 7, Table 12.2B) in **B** is based on a fugacity of $\text{CO}_2 = 3 \times 10^{-4}$ atm. The solubility limits of amorphous silica and calcite restrict the compositions of natural solutions to the lower left corner of **B**. Point *P* in that area represents an environment in which anorthite alters to gibbsite, a reaction that can cause the composition of the water to change. When the reaction path reaches line 5 (Table 12.2B), calcite begins to precipitate and the path turns to the right but does not advance beyond the gibbsite-kaolinite border (equation 2, Table 12.2B). Because of the way in which the solubility limit of calcite restricts the $[\text{Ca}^{2+}]/[\text{H}^+]^2$ ratio of water, authigenic anorthite cannot form by reactions of natural solutions with gibbsite, kaolinite, or pyrophyllite at 25°C.

Table 12.2 Stability Relations for Albite and Anorthite in Contact with Aqueous Solutions at 25°C and 1 atm pressure^a**A. Albite**

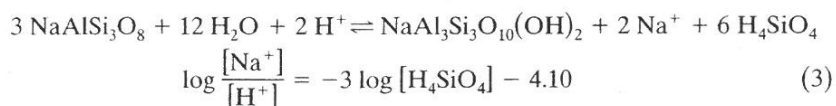
Albite-kaolinite



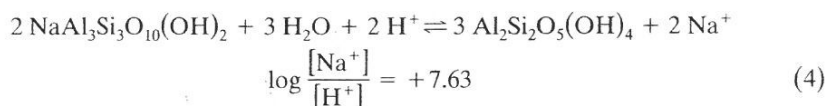
Kaolinite-gibbsite



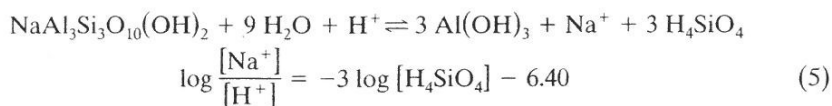
Albite-paragonite



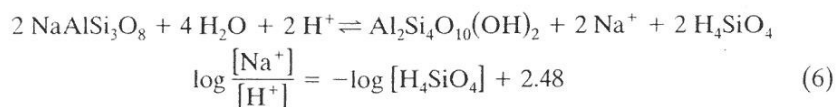
Paragonite-kaolinite



Paragonite-gibbsite



Albite-pyrophyllite



Pyrophyllite-Kaolinite

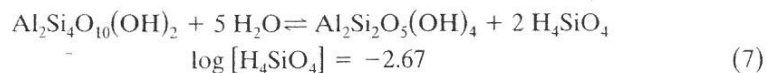
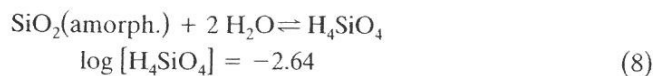
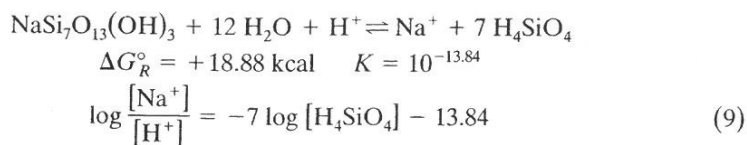


Table 12.2 (continued)

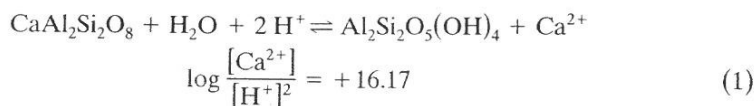
Solubility limit of amorphous silica



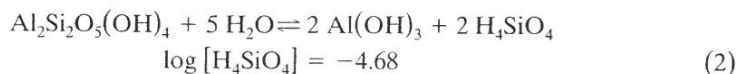
Solubility limit of magadiite

**B. Anorthite**

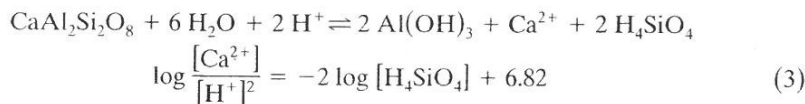
Anorthite-kaolinite



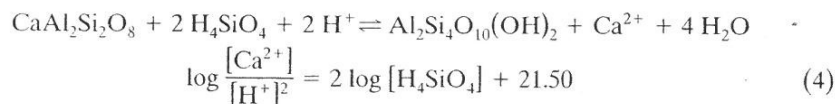
Kaolinite-gibbsite



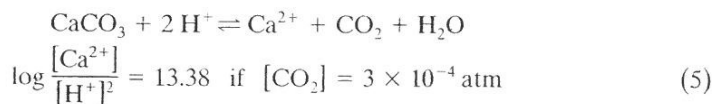
Anorthite-gibbsite



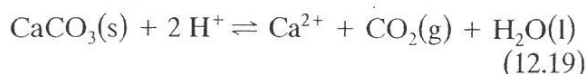
Anorthite-pyrophyllite



Solubility limit of calcite

^aBased on the thermodynamic data in Appendix B.

thereby avoid having to introduce the bicarbonate ion into the equations:



In the standard state, $\Delta G_R^\circ = -13.441$ kcal and hence $K = 10^{9.85}$. If $[\text{CO}_2] = 3 \times 10^{-4}$ atm:

$$\frac{[\text{Ca}^{2+}]}{[\text{H}^+]^2} = \frac{10^{9.85}}{3 \times 10^{-4}} = 10^{13.38} \quad (12.20)$$

and $\log [\text{Ca}^{2+}]/[\text{H}^+]^2 = 13.38$. Therefore, calcite precipitates when $\log [\text{Ca}^{2+}]/[\text{H}^+]^2 > 13.38$. Consequently, the decomposition of anorthite can result in the formation of gibbsite + calcite, kaolinite + calcite, or kaolinite + calcite + amorphous silica, depending on the environment to which anorthite is exposed. For example, if anorthite is placed into an environment represented by point *P* in the stability field of gibbsite, anorthite is converted to gibbsite in accordance with reaction 3 in Table 12.2B. If the system is closed and the water/rock ratio is small, the composition of the water changes because Ca^{2+} and H_4SiO_4 are released. If the fugacity of CO_2 is 3×10^{-4} atm, calcite begins to precipitate when the reaction path reaches the solubility boundary of calcite. As a result, the reaction path moves along the solubility boundary of calcite (line 5 in Figure 12.2B) because the activity of H_4SiO_4 increases as anorthite continues to be converted to gibbsite.

When the reaction path reaches the stability boundary between gibbsite and kaolinite (line 2 in Figure 12.2B), the remaining anorthite reacts to form kaolinite by reaction 1 in Table 12.2B. This reaction tends to increase the $[\text{Ca}^{2+}]/[\text{H}^+]^2$ ratio of the solution and therefore causes more calcite to precipitate. The reaction path does not advance beyond the gibbsite-kaolinite boundary because the conversion of anorthite to kaolinite does not increase the activity of H_4SiO_4 of the solution and because the $[\text{Ca}^{2+}]/[\text{H}^+]^2$ activity ratio cannot exceed the solubility limit of calcite. Therefore, the solubility limits of calcite and amorphous silica in Figure 12.2B restrict the compositions of natural waters to the lower left corner of the diagram.

Another consequence of the calcite solubility limit is that neither gibbsite nor kaolinite can be converted into anorthite because $\log [\text{Ca}^{2+}]/[\text{H}^+]^2$ of the solution cannot rise above 13.38. Therefore, the reactions between acidified water and anorthite that produce either kaolinite or gibbsite are not reversible at 25°C. If the fugacity of CO_2 is increased by a factor of 10 to 3×10^{-3} atm, $\log [\text{Ca}^{2+}]/[\text{H}^+]^2 = 12.38$ (equation 12.20), which causes calcite to precipitate at even lower values of the $[\text{Ca}^{2+}]/[\text{H}^+]^2$ ratio. Therefore, anorthite is *not* formed from clay minerals or aluminum hydroxide at Earth-surface temperatures, whereas albite and K-feldspar *do* form because their carbonates, oxides, and hydroxides are much more soluble than those of Ca or Mg and therefore do not impose limits on the activities of their ions in aqueous solutions.

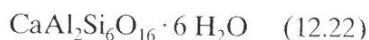
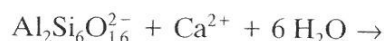
12.2 Formation of Zeolites

The zeolites are a large and highly diversified group of aluminosilicates that occur in certain igneous, metamorphic, and sedimentary rocks (Gottardi and Galli, 1985). They can form by direct precipitation from aqueous solution in vesicles and fractures of lava flow. However, zeolites also form in sedimentary rocks by reactions between volcanic glass, feldspars, feldspathoids, and other silicate minerals with saline solutions of high pH (Lisitzina and Butuzova, 1982; Petzing and Chester, 1979). The zeolites contain water of hydration, which is gradually forced out by increases in pressure and temperature. Hence, zeolites have been used to determine the depth of burial of lava flows and of volcanoclastic sedimentary rocks (Coombs et al., 1959; Walker, 1960; Sutherland, 1977). The crystal structures of zeolites contain channels and cavities that make these minerals useful as molecular sieves. Therefore, zeolites are used in the purification of water and other liquids and gases, for chemical separations, and for decontamination of radioactive waste (Barrer, 1978). Because of their many industrial uses, zeolites are synthesized for specific applica-

tions, but natural materials recovered from large sedimentary deposits are also used (Hay, 1966).

Zeolites are tektosilicates (Section 8.3) composed of silica tetrahedra linked at the corners by sharing oxygen atoms so that the ratio of Si to O is 1:2. The replacement of some Si atoms by Al atoms creates a *deficiency of positive charge* in the lattice because Al is trivalent whereas Si is tetravalent. The resulting *excess negative charge* is neutralized by incorporation of Na^+ or Ca^{2+} and, more rarely, of K^+ , Ba^{2+} , and Sr^{2+} ions. The zeolites have a more open lattice than other tektosilicates, such as the feldspars and feldspathoids, which allows them to contain water and makes them useful as molecular sieves. Some zeolites (chabazite) can be dehydrated by heating without permanently damaging the crystal lattice. Dehydrated chabazite can reabsorb the same amount of water it contained originally. Other zeolites (phillipsite and gismondine) experience major structural changes during dehydration below 200°C .

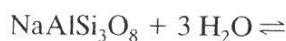
The chemical compositions of zeolites range widely because of variations in the number of Si atoms that are replaced by Al atoms, because the resulting charge deficiency is neutralized by varying proportions of Na^+ , K^+ , Ca^{2+} , Ba^{2+} , and Sr^{2+} , and because the amount of water may vary depending on pressure and temperature. Nevertheless, the chemical formulas of zeolites can be rationalized by normalizing the Al to two atoms, regardless of the numbers of SiO_2 "units" or water molecules. For example, the formula of stilbite ($\text{CaAl}_2\text{Si}_6\text{O}_{16} \cdot 6\text{H}_2\text{O}$) can be understood in terms of coupled substitution:



The composition of zeolites in sedimentary rocks may change continuously in response to changes in the physical and chemical state of their environment. The complexity of the chemical compositions of zeolites and their sensitivity to environmental change make it difficult to determine their standard free energies of formation (Johnson et al., 1982; Hemingway and Robie,

1984). Hence, thermodynamic constants of only a few zeolites (analcime, wairakite, laumontite, and leonhardite) are listed in Appendix B.

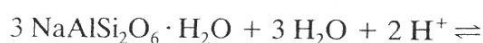
We begin by modifying the stability diagram of albite (Figure 12.2A) by including the zeolite analcime ($\text{NaAlSi}_2\text{O}_6 \cdot \text{H}_2\text{O}$). The reaction by which albite is transformed into analcime is easily constructed:



$$\Delta G_R^\circ = +5.741 \text{ kcal} \quad K = 10^{-4.21}$$

$$\log [\text{H}_4\text{SiO}_4] = -4.21 \quad (12.24)$$

When equation 12.24 is plotted as line 1 in Figure 12.3A, it becomes apparent that analcime must also react to form paragonite. Therefore, we construct the equation to represent this reaction:

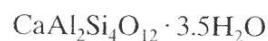
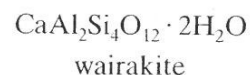


$$\Delta G_R^\circ = -6.045 \text{ kcal} \quad K = 10^{+4.43}$$

$$\log \frac{[\text{Na}^+]}{[\text{H}^+]} = -1.5 \log [\text{H}_4\text{SiO}_4] + 2.22 \quad (12.26)$$

The remainder of Figure 12.3A is identical to Figure 12.2A. We see that analcime is stable in environments having high activities of Na^+ but low activities of H^+ , which is typical of saline brines formed by evaporative concentration of surface water. Therefore, analcime commonly occurs in playa-lake deposits and in soils and volcaniclastic sediment in arid regions of the world. It occurs in very small crystals ranging from 5 to $100 \mu\text{m}$ but can be detected by x-ray diffraction (Hay, 1966).

Next, we consider the Ca-zeolites:



leonhardite



laumontite

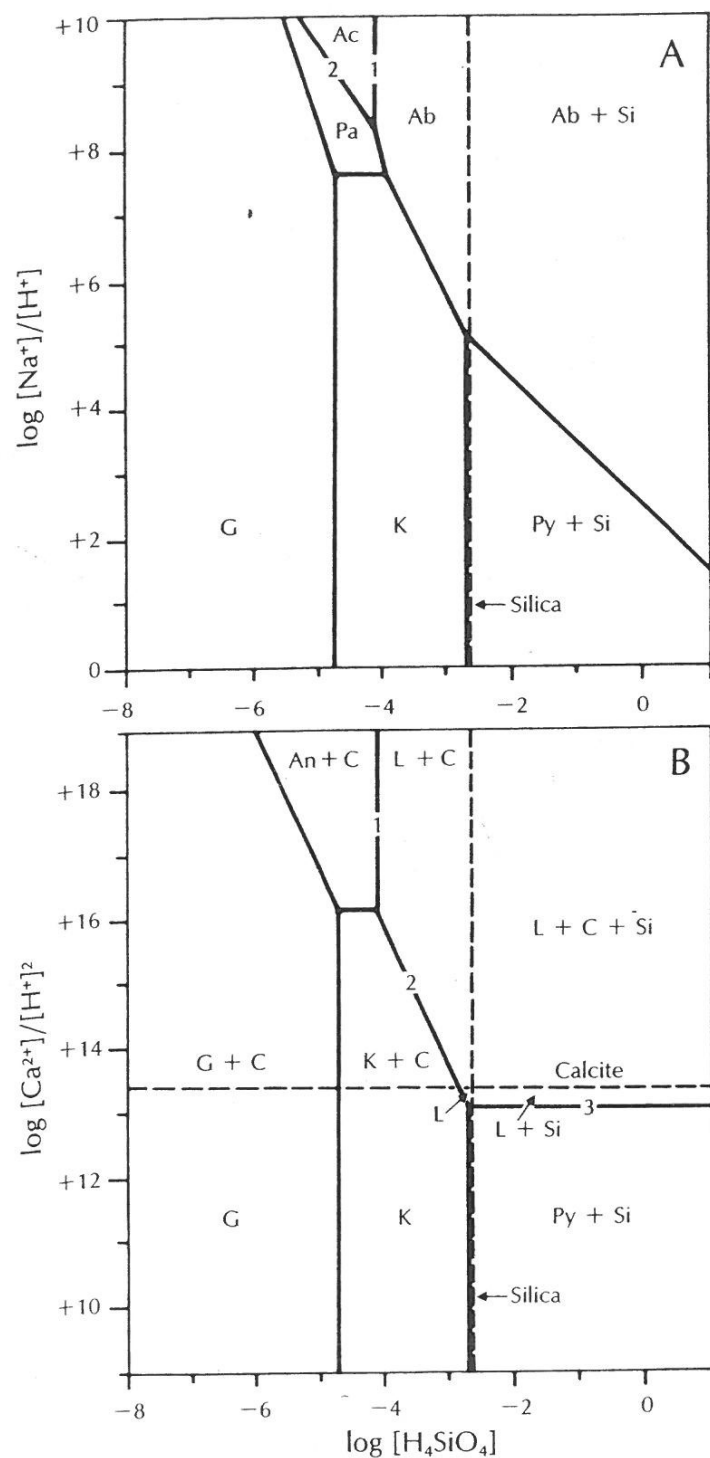
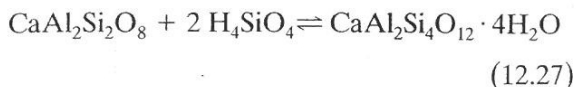


Figure 12.3 Stability of albite (A) and anorthite (B) with respect to analcime and laumontite, respectively, in the presence of acidified water at 25°C at 1 atm pressure. Ab = albite, Ac = analcime, Pa = paragonite, K = kaolinite, G = gibbsite, Si = amorphous silica, An = anorthite, L = laumontite. The equations for the numbered boundaries are given in the text and the others are the same as in Figure 12.2. Analcime can form when any of the solids in Figure 12.3A reacts with Na-rich brines of high pH. However, the brine compositions required to form laumontite at 25°C are restricted by the solubility limits of calcite and amorphous silica to the small area in Figure 12.3B labeled L.

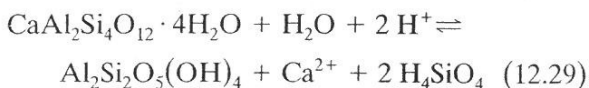
all of which have identical molar Si/Al ratios of 2:1 but differ in the number of water molecules they contain. Therefore, we choose laumontite to represent these minerals on the stability diagram in Figure 12.3B because it forms in near-surface environments and may subsequently convert to leonhardite or wairakite. Leonhardite forms by dehydration of laumontite upon exposure to the atmosphere, whereas wairakite becomes stable between 150 and 300°C at 1 atm pressure (Bowers et al., 1984). The transformation of anorthite to laumontite proceeds as follows (line 1, Figure 12.3B):



$$\Delta G_R^\circ = -11.32 \text{ kcal} \quad K = 10^{8.30}$$

$$\log [\text{H}_4\text{SiO}_4] = -4.15 \quad (12.28)$$

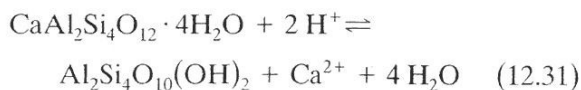
The decomposition of laumontite to kaolinite is represented by the equation:



$$\Delta G_R^\circ = -10.73 \text{ kcal} \quad K = 10^{7.87}$$

$$\log \frac{[\text{Ca}^{2+}]}{[\text{H}^+]^2} = -2 \log [\text{H}_4\text{SiO}_4] + 7.87 \quad (12.30)$$

Equation 12.30 has been plotted as line 2 in Figure 12.3B. Finally, the laumontite-pyrophyllite equilibrium is represented by:



$$\Delta G_R^\circ = -18.00 \text{ kcal} \quad K = 10^{+13.20}$$

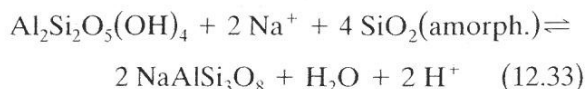
$$\log \frac{[\text{Ca}^{2+}]}{[\text{H}^+]^2} = +13.20 \quad (12.32)$$

Equation 12.32 forms line 3 in Figure 12.3B. All other boundaries are the same as in Figure 12.2B.

We see that laumontite has a large stability field in Figure 12.3B, but only a very small part of it lies within the compositional range of natural solutions. Therefore, anorthite, or any other Al-bearing compound, can alter to laumontite at

25°C only in those environments that lie within the small triangular field labeled "L." In addition, the conversion of anorthite to laumontite is irreversible because the calcite solubility boundary at atmospheric values of the fugacity of CO_2 prevents the $[\text{Ca}^{2+}]/[\text{H}^+]^2$ ratio of natural solutions from increasing to values within the anorthite field.

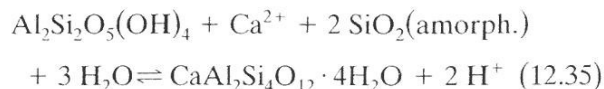
We can combine the stability diagrams of analcime and laumontite (Figure 12.3) by specifying that the solutions are saturated with respect to amorphous silica. This presents no problem for laumontite, which can be in equilibrium with such a solution. However, analcime, paragonite, and gibbsite are not stable in contact with silica-saturated solutions. Therefore, we anticipate that kaolinite is the stable phase in silica-saturated solutions having low activities of Na^+ and Ca^{2+} and that kaolinite can coexist in equilibrium with albite and laumontite as the activities of Na^+ and Ca^{2+} in the solution increase. Actually, pyrophyllite forms from kaolinite just below the silica-saturation limit (Figure 12.2). However, we prefer kaolinite because it is far more abundant as a weathering product of aluminosilicate minerals than pyrophyllite. Accordingly, we begin the construction of this diagram with the conversion of kaolinite to albite in a silica-saturated solution:



$$\Delta G_R^\circ = +13.899 \text{ kcal} \quad K = 10^{-10.19}$$

$$\log \frac{[\text{Na}^+]}{[\text{H}^+]} = +5.09 \quad (\text{line 1, Figure 12.4}) \quad (12.34)$$

The reaction of kaolinite with Ca^{2+} to form laumontite is represented by:



$$\Delta G_R^\circ = +17.941 \text{ kcal} \quad K = 10^{-13.15}$$

$$\log \frac{[\text{Ca}^{2+}]}{[\text{H}^+]^2} = +13.15 \quad (\text{line 2, Figure 12.4}) \quad (12.36)$$

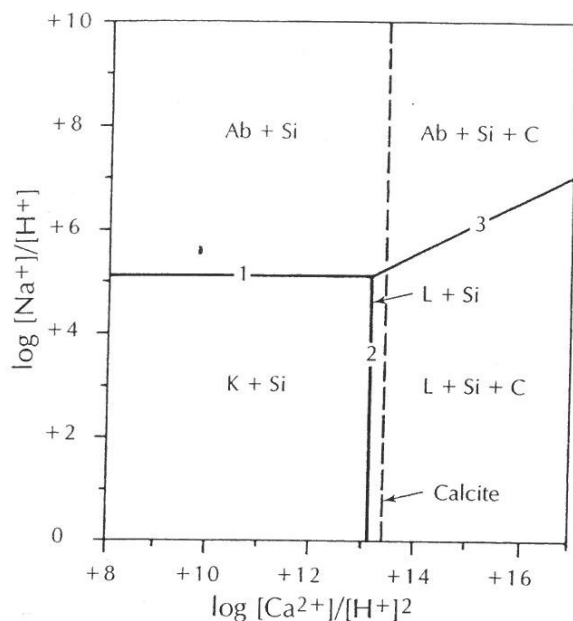
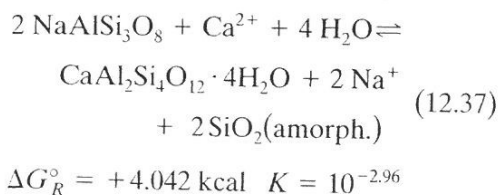


Figure 12.4 Stability diagram for albite (Ab), kaolinite (K), laumontite (L), and calcite (C) in contact with solutions saturated with respect to amorphous SiO_2 (Si) at 25°C . Analcime, paragonite, anorthite, and gibbsite were excluded from this diagram because they are unstable in silica-saturated solutions as shown in Figure 12.3. However, pyrophyllite is probably stable, but was excluded because kaolinite is more abundant as a weathering product of Al-silicates than pyrophyllite. The calcite-saturation limit is based on $[\text{CO}_2] = 3 \times 10^{-4}$ atm (Table 12.2B, equation 5). The numbered lines (1, 2, and 3) were derived from equations 12.33, 12.35, and 12.37, respectively.

Equilibrium between albite and laumontite in silica-saturated solutions is based on:



We note that reaction 12.37 is independent of the pH and therefore must be modified before it can be plotted in the coordinates of Figure 12.4.

Applying the Law of Mass Action to equation 12.37 yields:

$$\frac{[\text{Na}^+]^2}{[\text{Ca}^{2+}]} = 10^{-2.96} \quad (12.38)$$

We now divide the $[\text{Na}^+]^2$ and $[\text{Ca}^{2+}]$ in equation 12.38 by $[\text{H}^+]^2$:

$$\frac{[\text{Na}^+]^2/[\text{H}^+]^2}{[\text{Ca}^{2+}]/[\text{H}^+]^2} = 10^{-2.96} \quad (12.39)$$

Taking logarithms of equation 12.39 yields the desired equation:

$$2 \log \frac{[\text{Na}^+]}{[\text{H}^+]} - \log \frac{[\text{Ca}^{2+}]}{[\text{H}^+]^2} = -2.96 \quad (12.40)$$

and hence:

$$\log \frac{[\text{Na}^+]}{[\text{H}^+]} = 0.5 \log \frac{[\text{Ca}^{2+}]}{[\text{H}^+]^2} - 1.48 \quad (\text{line 3, Figure 12.4}) \quad (12.41)$$

The three lines derived from reactions 12.33, 12.35, and 12.37 have been plotted in Figure 12.4 together with the solubility limit of calcite (Table 12.2B, equation 5) for $[\text{CO}_2] = 3 \times 10^{-4}$ atm. The resulting diagram delineates the conditions under which albite decomposes to kaolinite + silica or laumontite + silica. Its virtue lies in the fact that the environmental conditions are described in terms of both Na^+ and Ca^{2+} , but it is limited by the instability of certain solids in the presence of silica-saturated solutions.

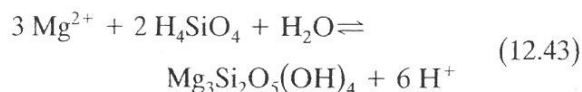
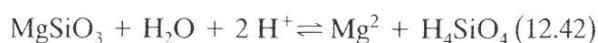
12.3 Magnesium Silicates

Magnesium forms a variety of silicate minerals, some of which include Al (phlogopite, cordierite, vermiculite, chlorite, and Mg-smectite), whereas others lack Al (forsterite, enstatite, serpentine, talc, and sepiolite). Some Mg silicates contain other cations, such as K^+ (phlogopite), Fe^{2+} , Ca^{2+} , and K^+ (vermiculite), and Fe^{2+} (Mg-montmorillonite). Therefore, these minerals can be shown only on diagrams that assume a specified activity of silicic

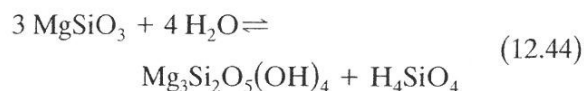
acid or require other assumptions. For this reason, Figure 12.5A includes the stability fields only of chlorite (Ct) together with pyrophyllite (Py), kaolinite (K), gibbsite (G), amorphous silica (Si), and magnesite (Ms). Brucite and periclase have solubility limits of $\log [\text{Mg}^{2+}]/[\text{H}^+]^2 = 16.71$ and $\log [\text{Mg}^{2+}]/[\text{H}^+]^2 = 21.59$, respectively, and are unstable in Figure 12.5A. Therefore, magnesite effectively limits the $[\text{Mg}^{2+}]/[\text{H}^+]^2$ ratio of natural solutions to values less than $10^{13.52}$, assuming that $[\text{CO}_2] = 3 \times 10^{-4}$ atm. The equations for these reactions are listed in Table 12.3A.

The Al-bearing silicates of Mg dissolve *incongruently* because gibbsite, kaolinite, and pyrophyllite are highly insoluble. This situation does not occur when the Mg silicates that lack Al dissolve in acidified water. These minerals therefore dissolve *congruently* and form Mg^{2+} and H_4SiO_4 . Consequently, the stabilities of the Mg-silicates are limited by the activities of Mg^{2+} , H^+ , and H_4SiO_4 of the solution with which they come in contact. The equations defining the solubilities of the Mg silicates in coordinates of $\log [\text{Mg}^{2+}]/[\text{H}^+]^2$ and $\log [\text{H}_4\text{SiO}_4]$ are listed in Table 12.3B and have been plotted as dashed lines in Figure 12.5B.

The solubilities of magnesite, serpentine, sepiolite, and amorphous silica outline a region in Figure 12.5B in which all of the minerals are *soluble* and are therefore *unstable*. When the activity ratio $[\text{Mg}^{2+}]/[\text{H}^+]^2$ and/or $[\text{H}_4\text{SiO}_4]$ of solutions in that region increase, either magnesite, serpentine, sepiolite, or amorphous silica precipitates, depending on which solubility line is contacted. The composition of a solution may change either because a soluble Mg silicate such as forsterite, talc, enstatite, or anthophyllite is dissolving in the water in a closed system with a small water/rock ratio, or because of evaporative concentration of surface water, or both. For example, if enstatite dissolves in acidified water represented by the point *P* in Figure 12.5B, the reaction path intersects the serpentine solubility line. Therefore, serpentine precipitates from the solution as enstatite dissolves. In effect, serpentine *replaces* enstatite by the reactions:



If Mg^{2+} is conserved by multiplying equation 12.42 by 3 before adding it to equation 12.43, the net result is:



Therefore, the conversion of enstatite to serpentine increases the activity of silicic acid. If enstatite continues to dissolve, the activity of silicic acid may increase until the solubility limit of sepiolite is reached. Continued solution of enstatite causes sepiolite to precipitate, and so on, until the solution becomes saturated with amorphous silica. However, the suggested sequence of events depends on assumptions about the reactions (i.e., Mg is conserved, enstatite continues to dissolve, no other reactions are taking place) and about the environment (closed, with small water/rock ratio, stable with respect to erosion, constant temperature) that are rarely satisfied in nature.

The replacement of enstatite by serpentine in reaction 12.44 can, in principle, achieve equilibrium at a specific value of $\log [\text{H}_4\text{SiO}_4]$. Since $\Delta G_R^\circ = -2.63$ kcal and $K = 10^{1.93}$, equilibrium between enstatite and serpentine occurs at $\log [\text{H}_4\text{SiO}_4] = 1.93$. This value *exceeds* the solubility limit of amorphous silica, which means that enstatite and serpentine cannot coexist in equilibrium in contact with natural solutions at 25°C. This is true also of enstatite–talc, serpentine–talc, forsterite–talc, serpentine–anthophyllite, and anthophyllite–talc because, in each case, the activity of H_4SiO_4 required for equilibrium is *greater* than that of a saturated solution of amorphous silica. Similarly, forsterite cannot coexist in equilibrium with enstatite, sepiolite, or anthophyllite because the $[\text{Mg}^{2+}]/[\text{H}^+]^2$ ratios required for equilibrium in each case exceed the solubility limit of magnesite.

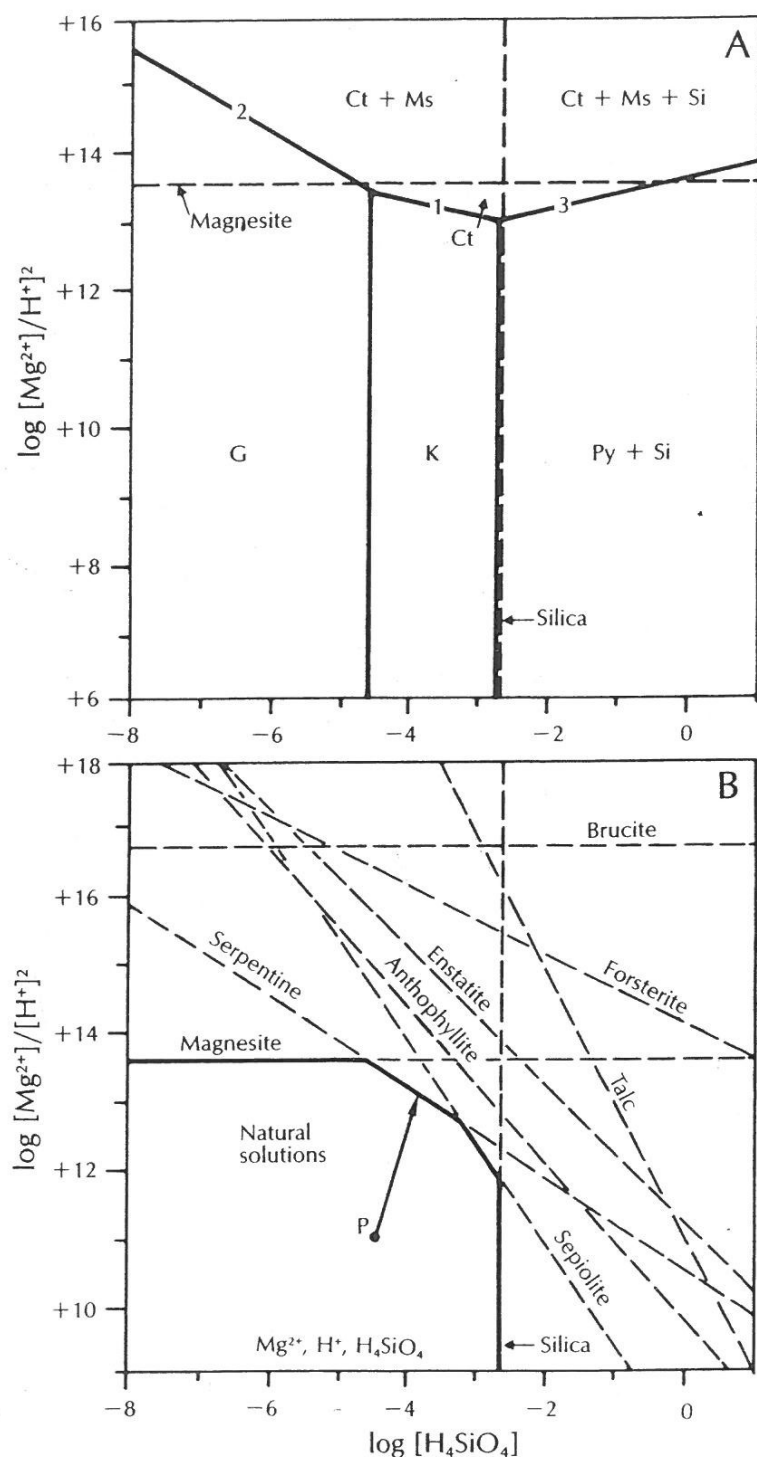
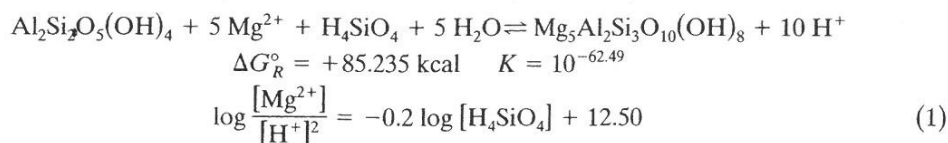


Figure 12.5 A: Stability diagram of Mg chlorite (Ct) (aluminosilicate) with respect to gibbsite (G), kaolinite (K), pyrophyllite (Py), amorphous silica (Si), and magnesite (Ms) at $[CO_2] = 3 \times 10^{-4}$ atm in the presence of acidified water at 25°C.

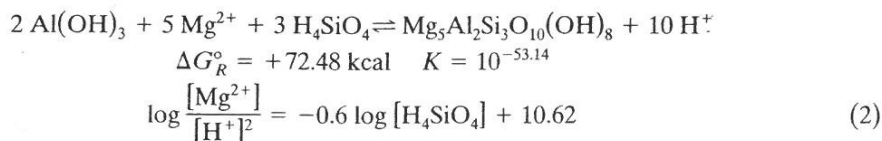
B: Solubility limits of pure Mg-silicates in acidified water at 25°C, based on equations in Table 12.3. Each of the minerals shown here dissolves in solutions whose $[Mg^{2+}]/[H^+]^2$ ratios and H_4SiO_4 activities are less than those of their respective solubility lines. Conversely, when these parameters increase in a solution in the lower left corner of the diagram, magnesite, serpentine, sepiolite, or amorphous silica precipitates, depending on which solubility line is reached by the solution. Consequently, these are the only stable minerals in contact with natural solutions on the surface of the Earth. The combination of solution and reprecipitation causes the replacement of forsterite, enstatite, talc, and anthophyllite by magnesite, serpentine, sepiolite, or amorphous silica. If amorphous silica precipitates first, and if one of the more soluble Mg silicates continues to dissolve, the $[Mg^{2+}]/[H^+]^2$ of the solution rises until sepiolite coprecipitates with silica. Similarly, the initial precipitation of magnesite may eventually result in the formation of serpentine, provided the activity of H_4SiO_4 increases sufficiently.

Table 12.3 Stability Relations among Mg Silicates in the Presence of Acidified Water at 25°C and 1 atm pressure^a**A. Mg-Al Silicates**

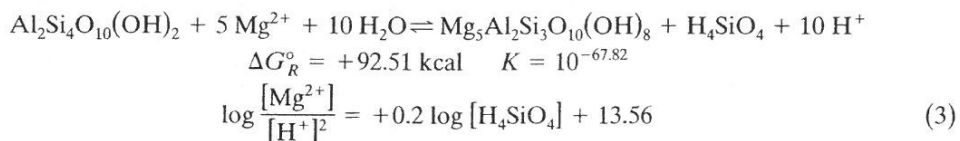
Kaolinite-chlorite



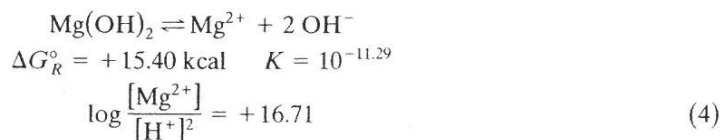
Gibbsite-chlorite



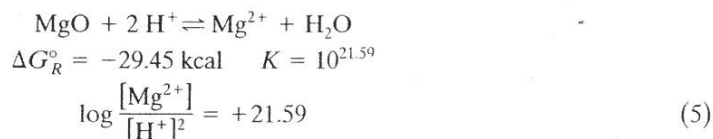
Pyrophyllite-chlorite



Solubility limit of brucite



Solubility limit of periclase



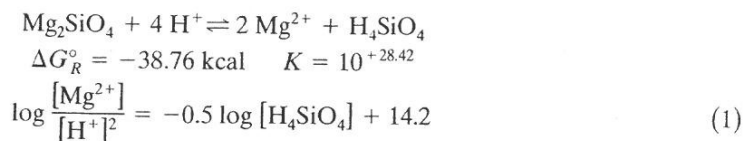
We conclude that the Mg-silicates whose solubilities are shown as *dashed* lines in Figure 12.5B are *unstable* in the presence of acidified water and therefore dissolve in it. As a result, the $[\text{Mg}^{2+}]/[\text{H}^+]^2$ and $[\text{H}_4\text{SiO}_4]$ values of

the solutions may rise until they reach the solubility limits of magnesite, serpentine, sepiolite, or amorphous silica, which are the only stable phases that can coexist with natural solutions at 25°C.

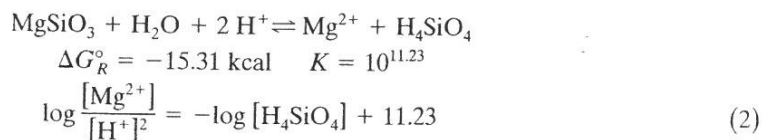
Table 12.3 (continued)

B. Solubility of Mg Silicates

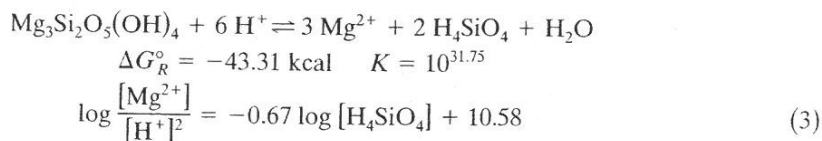
Forsterite



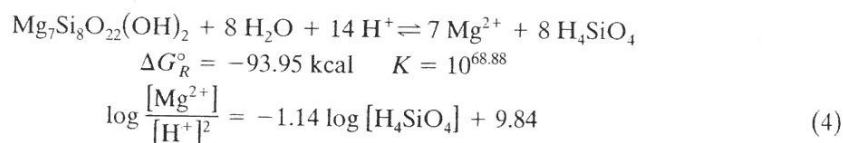
Enstatite



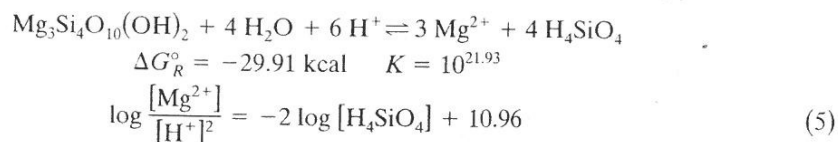
Serpentine



Anthophyllite



Talc

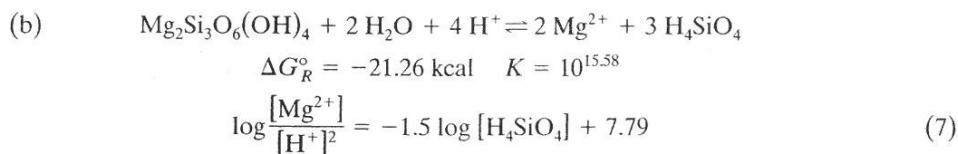
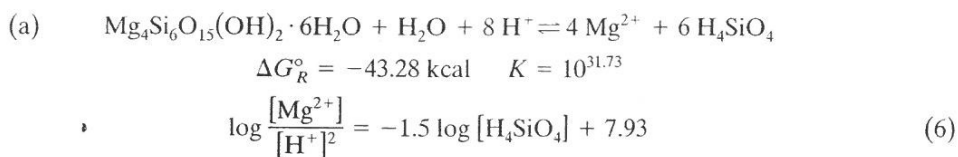
**12.4 Solubility Diagrams**

All minerals dissolve *congruently* when placed in *pure* water. For example, microcline dissolves congruently in pure water until the activities of

Al^{3+} ions in the solution increase sufficiently to stabilize gibbsite. Subsequently, microcline dissolves *incongruently* to form gibbsite plus K^+ and H_4SiO_4 . We therefore need to know how gibbsite dissolves and under what conditions it is stable

Table 12.3 (continued)

Sepiolite (see Appendix B)

^aBased on the thermodynamic data in Appendix B.

with respect to the different Al ions that can form in an aqueous solution.

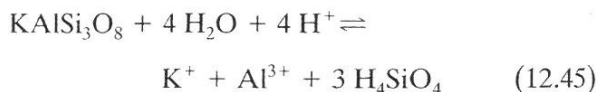
a. Gibbsite

Since gibbsite is the insoluble hydroxide of Al, we recall the discussion of the dissociation of bases and their amphoteric character from Section 9.4. At that stage in our discussion we were restricted to known dissociation constants, whereas we can now calculate them from values of ΔG_R° for any reaction we wish to study. The Appendix B contains G_f° values for Al^{3+} , $\text{Al}(\text{OH})^{2+}$, $\text{Al}(\text{OH})_2^+$, $\text{Al}(\text{OH})_3^0$, $\text{Al}(\text{OH})_4^-$, and $\text{Al}(\text{OH})_5^{2-}$, all of which can form when an Al-bearing mineral dissolves congruently at different pH values. The reactions by which these ions form from gibbsite and their equilibrium constants at 25°C are listed in Table 12.4.

The equations derived in this way outline the stability field of gibbsite in Figure 12.6. Outside this area, gibbsite dissolves congruently to form ions whose relative abundances become equal at certain pH values corresponding to the points of intersection of the solubility lines. For example, Al^{3+} is the dominant ion at $\text{pH} < 4.65$, and $\text{Al}(\text{OH})_2^+$ dominates at pH values from 4.65 to 5.7, as shown in Figure 12.6. The solubility of gibbsite at a specified pH is the sum of the concentrations of all Al-bearing species that exist in a saturated solution in equilibrium with gibbsite.

The solubility of this mineral has a minimum at pH from 6 to 8 and rises at $\text{pH} > 8$ and < 6 in accordance with the amphoteric character of Al.

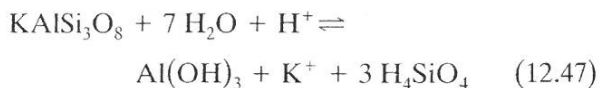
When microcline dissolves congruently in very dilute solutions having a $\text{pH} < 4.65$, it releases K^+ , Al^{3+} , and H_4SiO_4 according to the reaction:



When microcline dissolves congruently in water of composition *P* in Figure 12.6, both the activity of Al^{3+} and the pH rise along the reaction path. If the system is closed and its water/rock ratio is small, the solution may become saturated with respect to gibbsite, which then precipitates:



The net result, obtained by adding equations 12.45 and 12.46, is that microcline then dissolves *incongruently* to form gibbsite plus K^+ and H_4SiO_4 :



If microcline continues to dissolve, the composition of the solution moves across the gibbsite stability field in Figure 12.1 until it reaches the gibbsite-kaolinite boundary (line 2). At that

Chapter 14, where we review some basic principles of electrochemistry and apply them to the construction of the so-called Eh–pH diagrams.

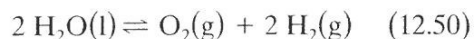
In Section 14.5 we take up the construction of fugacity diagrams based on equilibria between solid compounds and certain gases, including O_2 , CO_2 , and S_2 . The solid–gas equilibria of both Fe and Cu include possible *changes* in their *oxidation states*, but the reactions are quite uncomplicated. Thus we can present fugacity diagrams here rather than in Chapter 14.

12.5 Fugacity Diagrams

The oxides, carbonates, and sulfides of metals can be considered to be in equilibrium with gaseous O_2 , CO_2 , and S_2 , respectively. For example, the fugacity of O_2 ranges from about 2×10^{-1} atm in air at the surface of the Earth to much lower values in certain anoxic environments such as peat bogs, stagnant basins, and deep-sea trenches. Therefore, the first question we need to consider is the range of fugacities of O_2 that can occur on the surface of the Earth.

a. Stability Limits of Water

We know that liquid water dissociates into H^+ and OH^- ions and that the ion activity product of water at $25^\circ C$ is about 1×10^{-14} . However, water must also maintain an equilibrium with the gases O_2 and H_2 :



The standard free energy change of this reaction is $\Delta G_R^\circ = +113.374$ kcal and its equilibrium constant at $25^\circ C$ is $K = 10^{-83.1}$. Therefore, at equilibrium:

$$[O_2][H_2]^2 = 10^{-83.1} \quad (12.51)$$

The total pressure of all gases occurring naturally on the surface of the Earth must be 1 atm or less. If the pressure of gases rises to higher values, they expand against the atmosphere, which means that bubbles form in water exposed to the atmosphere

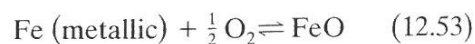
and the gases escape. Therefore, the fugacities of O_2 and H_2 must each remain <1.0 atm if liquid water is to be stable at sea level on the surface of the Earth. If the fugacity of H_2 alone is 1.0 atm, then the fugacity of O_2 must be as low as it can get in the presence of liquid water. Setting $[H_2] = 1.0$ in equation 12.51 yields:

$$[O_2] = 10^{-83.1} \text{ atm} \quad (12.52)$$

Therefore, the fugacity of O_2 in the presence of liquid water may vary from 1.0 to $10^{-83.1}$ atm. Similarly, the fugacity of H_2 can vary from 1.0 to $10^{-41.6}$ atm, based on equation 12.51.

b. Oxides and Carbonates of Iron

Metallic Fe exposed to the air is known to form the oxides wüstite (FeO), magnetite (Fe_3O_4), and hematite (Fe_2O_3). Therefore, we can represent the oxidation of metallic Fe by a series of equations starting with the formation of wüstite:

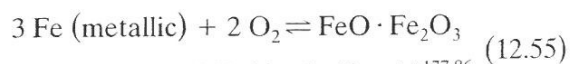


For this reaction, $\Delta G_R^\circ = -58.7$ kcal and $K = 10^{+43.0}$. Therefore, at equilibrium:

$$\frac{1}{[O_2]^{1/2}} = 10^{43.0} \quad (12.54)$$

and $[O_2] = 10^{-86.1}$. Evidently, wüstite can coexist in equilibrium with metallic Fe only when the fugacity of O_2 is less than is permitted in normal terrestrial environments.

We can continue either by considering the transformation of wüstite to magnetite, which is a mixed oxide ($FeO \cdot Fe_2O_3$), or by relating magnetite directly to metallic Fe. The latter is preferable because wüstite does not occur in terrestrial rocks. Therefore, we consider the reaction:

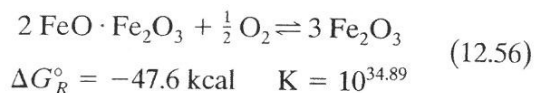


$$\Delta G_R^\circ = -242.6 \text{ kcal} \quad K = 10^{177.86}$$

It follows that at equilibrium the fugacity of O_2 must be $10^{-88.92}$ atm, which is also less than the lowest fugacity of O_2 in equilibrium with liquid

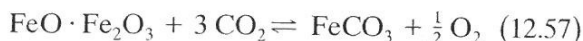
water. Therefore, metallic Fe cannot be stable on the surface of the Earth, where the fugacity of O_2 must be $>10^{-83.1}$ atm. However, magnetite *does* occur in surface environments and, therefore, is stable at O_2 fugacities that are compatible with liquid water.

We consider next the reaction between magnetite and hematite in the hope that this equilibrium will provide a stability limit for magnetite that is within the range of permissible O_2 fugacities:



Thus, magnetite and hematite are in equilibrium at $[O_2] = 10^{-69.8}$ atm, which is within the permissible range of O_2 fugacities. Hematite is the most O-rich oxide of Fe and can coexist with water at O_2 fugacities from $10^{-69.8}$ to 1.0 atm.

Next, we add siderite (FeCO_3) to the system by reacting magnetite and hematite with CO_2 gas. Since Fe in siderite is divalent, we start with magnetite because it also contains divalent Fe. We construct the desired equation with magnetite as the reactant and siderite as the product. Then we add CO_2 to the reactant side to balance the C in siderite. Finally, we complete the equation by adding O_2 as needed for balance:



Note that this is an oxidation-reduction reaction and that the equations *cannot* be balanced in terms of mass only, as we just did. In this case, two of the three divalent Fe ions in siderite give up one electron each for a total of two. These electrons are accepted by one atom of O_2 gas, which becomes one of the O^{2-} ions in the magnetite. Equation 12.57 is reversed because we have treated magnetite as the reactant rather than as the product. The standard free energy change of reaction 12.57 is $\Delta G_R^\circ = +44.48 \text{ kcal}$ and $k = 10^{-32.61}$. Therefore, at equilibrium:

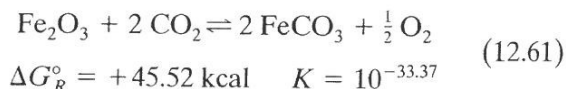
$$\frac{[O_2]^{1/2}}{[CO_2]^3} = 10^{-32.61} \quad (12.58)$$

Taking logarithms to the base 10:

$$\frac{1}{2} \log [O_2] - 3 \log [CO_2] = -32.61 \quad (12.59)$$

$$\log [O_2] = 6 \log [CO_2] - 65.22 \quad (12.60)$$

Similarly, hematite reacts with CO_2 to form siderite according to:



and:

$$\log [O_2] = 4 \log [CO_2] - 66.74 \quad (12.62)$$

The equations representing solid-gas equilibria derived above have been plotted in Figure 12.8 in coordinates of $\log [O_2]$ and $\log [CO_2]$. The result is a *fugacity diagram* for the oxides and carbonates of Fe that indicates the fugacities of the two gases that are required to stabilize each of the minerals we have considered. The diagram also predicts the direction of solid-gas reactions that occur when a particular mineral is placed in an environment in which another mineral is stable.

For example, so-called *ironstone concretions* composed of siderite may be exposed to the atmosphere as a result of erosion of the enclosing sedimentary rocks. Figure 12.8 indicates that hematite is stable in contact with the atmosphere. Therefore, siderite is transformed into hematite in accordance with equation 12.61. At equilibrium between siderite and hematite:

$$\frac{[O_2]^{1/2}}{[CO_2]^2} = 10^{-33.37} \quad (12.63)$$

However, in the atmosphere $[O_2] = 2 \times 10^{-1}$ atm and $[CO_2] = 3 \times 10^{-4}$ atm. Therefore, siderite can coexist in equilibrium with hematite in contact with the atmosphere only when:

$$[CO_2] = \left[\frac{[O_2]^{1/2}}{K} \right]^{1/2} = \left[\frac{(O_2)^{1/2}}{10^{-33.37}} \right]^{1/2}$$

$$= 10^{16.51} \text{ atm} \quad (12.64)$$

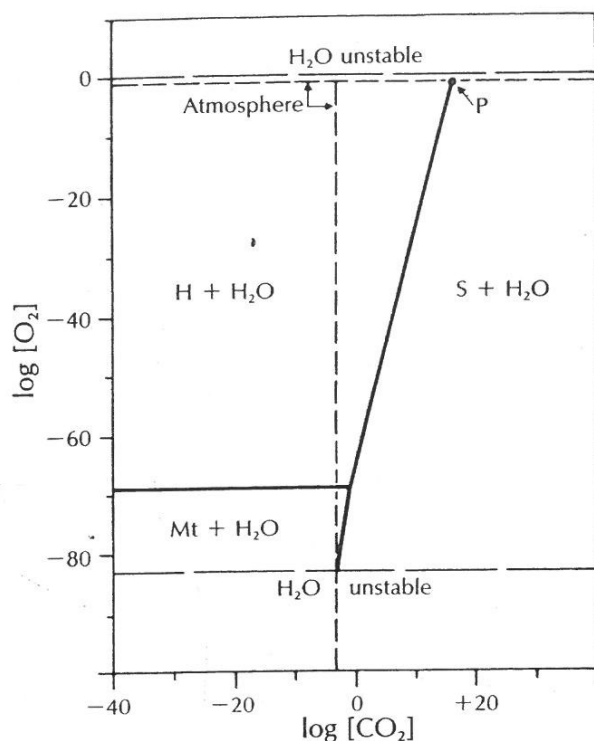


Figure 12.8 Fugacity diagram for the oxides and carbonate of Fe in terms of the fugacities of O_2 and CO_2 , where Mt = magnetite, H = hematite, and S = siderite. Note that siderite and magnetite are not stable in contact with the atmosphere, where $[O_2] = 2 \times 10^{-1}$ atm and $[CO_2] = 3 \times 10^{-4}$ atm. Therefore, when siderite or magnetite is exposed to the atmosphere, it converts to hematite. The diagram is based on equations 12.52 (stability limit of water) and 12.56, 12.58, and 12.61. Point P represents conditions when siderite is stable in contact with O_2 of the atmosphere.

which is identified by point P in the diagram. Since the fugacity of CO_2 in the atmosphere is only 3×10^{-4} atm, reaction 12.62 is not in equilibrium in contact with the atmosphere but runs to the left, thereby converting siderite to hematite. Therefore, ironstone concretions should weather to hematite upon exposure to the air. This prediction is confirmed by the presence of a hematite rind around ironstone concretions where the thickness of the rind depends on the rate of the reaction and the duration of the exposure.

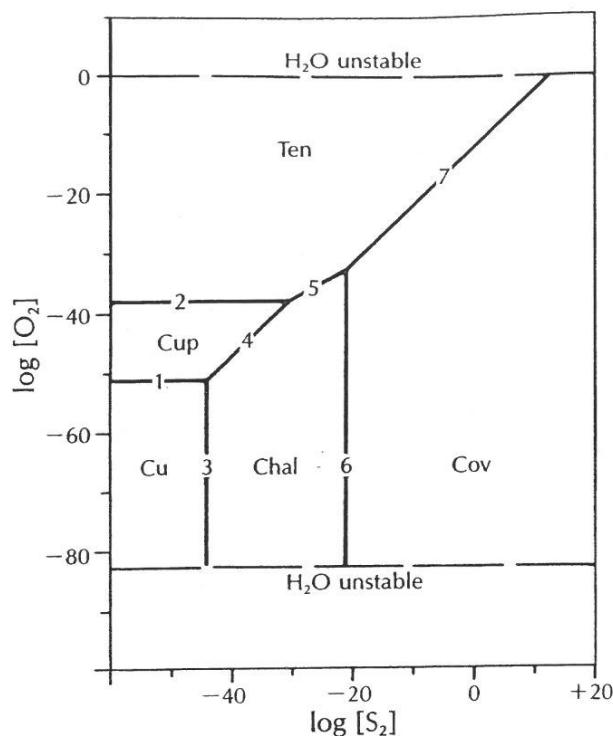


Figure 12.9 Fugacity diagram for the oxides and sulfides of Cu. Cu = metallic Cu, Cup = cuprite (Cu_2O), Ten = tenorite (CuO), Chal = chalcocite (Cu_2S), and Cov = covellite (CuS), based on equations in Table 12.5. Metallic Cu is stable in the presence of water, but when it is exposed to the atmosphere, it reacts to form tenorite. Note that G_f° of $S_2(g)$ is not zero because the stable form of S in the standard state is the rhombic solid.

c. Oxides and Sulfides of Copper

Copper forms the oxides cuprite (Cu_2O) and tenorite (CuO) and the sulfides chalcocite (Cu_2S) and covellite (CuS). Note that Cu is univalent in cuprite and chalcocite and divalent in tenorite and covellite. Therefore, when metallic Cu forms Cu^+ and when Cu^+ forms Cu^{2+} , electrons are transferred either to O_2 or to S_2 , depending on the compound that is produced. The equations required to construct the fugacity diagram for Cu in Figure 12.9 are listed in Table 12.5.

We see that native Cu is stable in the presence of water, whereas native Fe is not. However,

Table 12.5 Fugacity Diagram of Copper Oxides and Sulfides at 25°C^a

$2 \text{ Cu (metal)} + \frac{1}{2} \text{ O}_2 \rightleftharpoons \text{Cu}_2\text{O}$	
$\Delta G_R^\circ = -35.1 \text{ kcal}$	$K = 10^{25.73}$
$\log [\text{O}_2] = -51.5$	(1)
$\text{Cu}_2\text{O} + \frac{1}{2} \text{ O}_2 \rightleftharpoons 2 \text{ CuO}$	
$\Delta G_R^\circ = -26.3 \text{ kcal}$	$K = 10^{19.28}$
$\log [\text{O}_2] = -38.6$	(2)
$2 \text{ Cu (metal)} + \frac{1}{2} \text{ S}_2 \rightleftharpoons \text{Cu}_2\text{S}$	
$\Delta G_R^\circ = -30.08 \text{ kcal}$	$K = 10^{22.05}$
$\log [\text{S}_2] = -44.1$	(3)
$\text{Cu}_2\text{O} + \frac{1}{2} \text{ S}_2 \rightleftharpoons \text{Cu}_2\text{S} + \frac{1}{2} \text{ O}_2$	
$\Delta G_R^\circ = +5.02 \text{ kcal}$	$K = 10^{-3.60}$
$\log [\text{O}_2] = \log [\text{S}_2] - 7.4$	(4)
$2 \text{ CuO} + \frac{1}{2} \text{ S}_2 \rightleftharpoons \text{Cu}_2\text{S} + \text{O}_2$	
$\Delta G_R^\circ = +31.32 \text{ kcal}$	$K = 10^{-22.96}$
$\log [\text{O}_2] = 0.5 \log [\text{S}_2] - 23.0$	(5)
$\text{Cu}_2\text{S} + \frac{1}{2} \text{ S}_2 \rightleftharpoons 2 \text{ CuS}$	
$\Delta G_R^\circ = -14.48 \text{ kcal}$	$K = 10^{10.61}$
$\log [\text{S}_2] = -21.2$	(6)
$\text{CuO} + \frac{1}{2} \text{ S}_2 \rightleftharpoons \text{CuS} + \frac{1}{2} \text{ O}_2$	
$\Delta G_R^\circ = +8.42 \text{ kcal}$	$K = 10^{-6.17}$
$\log [\text{O}_2] = \log [\text{S}_2] - 12.34$	(7)

^aBased on the thermodynamic data in Appendix B.

when metallic Cu is exposed to O₂ at fugacities $>10^{-51.5}$ atm, it forms the oxide cuprite (Cu₂O) or tenorite (CuO), depending on the O₂ fugacity. Similarly, Cu reacts with S₂ to form the sulfides chalcocite (Cu₂S) and covellite (CuS). Copper also forms mixed sulfides with Fe including chalcopyrite (CuFeS₂), cubanite (CuFe₂S₃), and bornite (Cu₃FeS₄).

Fugacity diagrams of different metals can be superimposed on each other to form composite diagrams that indicate the stability limits of the compounds of both metals. Such a diagram has been constructed in Figure 12.10 for the oxides and carbonates of Fe and Cu. The diagram defines the fugacities of O₂ and CO₂ that permit certain assemblages of these minerals to coexist in equilibrium in the presence of water.

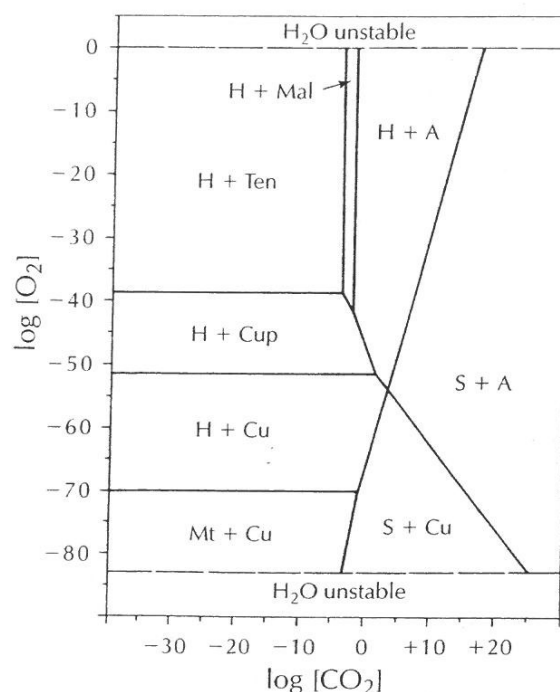


Figure 12.10 Combined fugacity diagram for Fe and Cu and their oxides and carbonates. H = hematite (Fe₂O₃); Mt = magnetite (Fe₃O₄); S = siderite (FeCO₃); Cu = native Cu; Cup = cuprite (Cu₂O); Ten = tenorite (CuO); Mal = malachite [CuCO₃ · Cu(OH)₂]; A = azurite [2CuCO₃ · Cu(OH)₂]. The boundaries between tenorite and malachite and between malachite and azurite were drawn after Garrels and Christ (1965). This diagram indicates the environmental conditions under which certain assemblages of Fe and Cu minerals can occur together.

For example, the diagram indicates that the carbonates of Cu (malachite and azurite) form at lower fugacities of CO₂ than siderite and therefore can coexist in equilibrium with hematite and tenorite. Consequently, malachite, azurite, tenorite, and hematite are common weathering products of Fe–Cu sulfide minerals and occur in the oxidized part of sulfide ore bodies (gossans), whereas siderite is unstable at the surface of the Earth (see also Section 21.2). Figure 12.10 also points out that native Cu has a large stability field in the absence of S and can coexist with magnetite,

hematite, and siderite, provided the fugacity of O_2 is low enough. When native Cu is exposed to the atmosphere, it *tarnishes* by forming a layer of tenorite or malachite or both, depending on the fugacity of CO_2 .

12.6 Summary

When the silicate minerals of igneous and metamorphic rocks are exposed to acidified water, they may dissolve *incongruently* to produce clay minerals and zeolites as well as insoluble oxides and hydroxides, depending on the environmental conditions. The chemical reactions that take place during these transformations can be used to define the conditions required for equilibrium when both reactants and products are stable. Therefore, equations can be derived from the Law of Mass Action, which specify the conditions required for equilibrium between the solids.

The resulting mineral stability diagrams define the environmental conditions in which each of the solids in a given compositional system is stable. In addition, the diagrams predict what happens to a mineral when it is exposed to conditions outside of its own stability field. The reactions that occur in such cases not only cause transformations among the solids but also affect the chemical composition, or *quality*, of water both above and below the surface of the Earth. When the reactions occur in closed systems with small water/rock ratios, the chemical composition of the water may evolve along curved reaction paths. Such conditions exist in the pore spaces or fractures of rocks in subsurface where reactions between minerals and water not only contribute to the chemical evolution of the water but also cause the deposition of mineral cements.

The solubility limits of certain compounds, including calcite, magnesite, amorphous silica, and others, which can precipitate from natural solutions, significantly restrict the chemical compositions of aqueous solutions in nature.

Therefore, certain minerals cannot be in equilibrium in the presence of water and also cannot form by transformation from other minerals. This phenomenon is exemplified by anorthite, which does not form as an authigenic mineral at Earth-surface temperatures, whereas albite and K-feldspar do occur in places where alkali-rich brines can react with volcanic glass, clay minerals, or other kinds of solids.

The conditions required for stability of minerals can also be expressed by their congruent solution or by their interaction with gases such as O_2 , CO_2 , and S_2 . For example, the non-aluminous Mg silicates dissolve congruently to form Mg^{2+} and H_4SiO_4 , which may precipitate as a different Mg silicate whose solubility is exceeded. The combination of congruent solution of a "primary" Mg silicate and the precipitation of a "secondary" mineral results in the replacement of one mineral by another. The apparent conversion of enstatite to serpentine is an example of such paired reactions.

The equilibrium between water and its constituent gases, O_2 and H_2 , defines the stability range of water within which all reactions on the surface of the Earth take place. Solid oxides, carbonates, and sulfides must similarly remain in equilibrium with their constituent gases O_2 , CO_2 , and S_2 , respectively. These solid-gas reactions can be used to construct fugacity diagrams that contain the stability fields of the solids appropriate to a particular system. The fugacity diagrams of different metals can be superimposed on each other to form combination diagrams. These diagrams effectively subdivide the stability field of water into environments in which suites of minerals of different elements can coexist in equilibrium.

The virtue of activity and fugacity diagrams is that they depict the relationships between minerals and aqueous solutions or gases in a readily usable form. As a result, idle speculation is replaced by understanding and progress is made in understanding the chemical reactions that contribute to geological processes.

Problems

1. Construct the equations necessary to introduce pyrophyllite into the activity diagram for microcline in Figure 12.1.
2. Construct the equations necessary to introduce Na-smectite¹ ($\text{Na}_{0.33}\text{Al}_{2.33}\text{Si}_{3.67}\text{O}_{10}(\text{OH})_2$, $G_f^\circ = -1277.76 \text{ kcal/mol}$) into the activity diagram for albite in Figure 12.2.
3. Construct a solubility diagram for goethite (FeOOH) as a function of pH.
4. What are the environmental conditions in terms of $[\text{Mg}^{2+}]/[\text{H}^+]^2$ and $[\text{H}_4\text{SiO}_4]$ at which enstatite and forsterite coexist in equilibrium?
5. Is equilibrium possible between enstatite and forsterite in the presence of aqueous solutions in nature?
6. What mineral precipitates first when forsterite dissolves congruently in water having the composition of point *P* in Figure 12.5?
7. Construct a fugacity diagram for the sulfides and carbonates of Pb.
8. Construct a fugacity diagram for the sulfides and carbonates of Zn.
9. Combine the two diagrams for Zn and Pb.
10. What are the fugacities of O_2 and CO_2 at which native Cu, cuprite, and azurite coexist in equilibrium?
11. Can hematite and siderite both join that assemblage in equilibrium with each other?

References

- BARRER, R. M., 1978. *Zeolites and Clay Minerals as Sorbents and Molecular Sieves*. Academic Press, London, 497 pp.
- BOWERS, T. S., K. J. JACKSON, and H. C. HELGESON, 1984. *Equilibrium Activity Diagrams*. Springer-Verlag, New York, 397 pp.
- COOMBS, D. S., A. J. ELLIS, W. S. FYFE, and A. M. TAYLOR, 1959. The zeolite facies, with comments on the interpretation of hydrothermal syntheses. *Geochim. Cosmochim. Acta*, **17**: 52–107.
- GARRELS, R. M., and C. L. CHRIST, 1965. *Solutions, Minerals and Equilibria*. Harper & Row, New York, 450 pp.
- GOTTARDI, G., and E. GALLI, 1985. *Natural Zeolites*, Springer-Verlag, New York, 390 pp.
- HAY, R. L., 1966. Zeolites and zeolitic reactions in sedimentary rocks. *Geol. Soc. Amer., Spec. Paper No. 85*, 130 pp.
- HELGESON, H. C., R. M. GARRELS, and F. T. MACKENZIE, 1969. Evaluation of irreversible reactions in geochemical processes involving minerals and aqueous solutions. II. Application. *Geochim. Cosmochim. Acta*, **33**: 455–481.
- HEMINGWAY, B. S., and R. A. ROBIE, 1984. Thermodynamic properties of zeolites: low-temperature heat capacities and thermodynamic functions for phillipsite and clinoptilolite. Estimates of the thermochemical properties of zeolitic water at low temperature. *Amer. Mineral.*, **69**: 692–700.
- JOHNSON, G. K., H. E. FLOTOW, P. A. G. O'HARE, and W. S. WISE, 1982. Thermodynamic studies of zeolites: analcime and dehydrated analcime. *Amer. Mineral.*, **67**: 736–748.
- KASTNER, M., 1971. Authigenic feldspars in carbonate rocks. *Amer. Mineral.*, **56**: 1413–1442.
- KASTNER, M., and R. SIEVER, 1979. Low temperature feldspars in sedimentary rocks. *Amer. J. Sci.*, **279**: 435–479.
- LISITZINA, N. A., and G. YU. BUTUZOVA, 1982. Authigenic zeolites in the sedimentary mantle of the world ocean. *Sed. Geol.*, **31**: 33–42.
- MENSING, T. M., and G. FAURE, 1983. Identification and age of neoformed Paleozoic feldspar (adularia) in a Precambrian basement core from Scioto County, Ohio, U.S.A. *Contrib. Mineral Petrol.*, **82**: 327–333.
- PETZING, J., and R. CHESTER, 1979. Authigenic marine zeolites and their relationship to global volcanism. *Marine Geol.*, **29**: 253–272.
- STEINMANN, P., P. C. LICHTNER, and W. SHOTYK, 1994. Reaction path approach to mineral weathering reactions. *Clays and Clay Minerals*, **42**: 197–206.
- SURDAM, R. C., and H. P. EUGSTER, 1976. Mineral reactions in the sedimentary deposits of the Lake Magadi region, Kenya. *Geol. Soc. Amer. Bull.*, **87**: 1739–1752.
- SUTHERLAND, F. L., 1977. Zeolite minerals in the Jurassic dolerites of Tasmania: Their use as possible indicators of burial depth. *Geol. Soc. Aust. J.*, **24**: 171–178.
- WALKER, G. P. L., 1960. Zeolite zones and dike distribution in relation to the structure of the basalts of eastern Iceland. *J. Geol.*, **68**: 515–528.

Clay Minerals

Clay minerals form one of the largest and most highly diversified groups of minerals known. The majority of clay minerals are aluminosilicates and they crystallize as phyllosilicates. However, some clay minerals are Mg-silicates or Fe-silicates, others are *not* crystalline, or they are not phyllosilicates. The property common to all clay minerals is the small size of their crystals or grains with diameters less than 2 micrometers (μm). As a result, sediment in that size range is sometimes referred to as the *clay-size fraction* or even as *clay*. This usage is incorrect in geochemistry because clay minerals are defined on the basis of their crystal structure and chemical composition and not by the size of their crystals or grains. Moreover, the clay-size fraction of sediment frequently includes small grains of other minerals (quartz, feldspar, calcite, etc.), as well as grains of insoluble oxides or hydroxides of Fe, Al, Mn, and so on, which may be partly or completely amorphous. Nevertheless, clay minerals are typically fine grained and therefore cannot be identified solely by their physical or optical properties. Instead, the classification of clay minerals relies heavily on their crystallographic and certain physical properties that are revealed by x-ray diffraction.

Clay mineralogy has evolved into an important field with applications in geology, agronomy, ceramic engineering, soil mechanics, and geochemistry. We will be concerned primarily with the classification of clay minerals, with their stability in different geochemical environments, with ion exchange phenomena, and with the use of clay minerals for dating sedimentary rocks. The presence of clay minerals in natural environments frequently makes a significant

difference to the geochemical processes that may be occurring. Since we encountered the clay minerals kaolinite, pyrophyllite, and smectite in Chapter 12, it is now necessary for us to examine these minerals in more detail.

13.1 Crystal Structure

Clay minerals are divided into three groups based on their crystallographic properties: (1) platy clay minerals (phyllosilicates), (2) fibrous clay minerals, and (3) amorphous clay. The platy clay minerals are much more abundant than the other two groups and include most of the minerals that traditionally make up the clay minerals. The fibrous clays *sepiolite* and *palygorskite* (also known as attapulgite), though not as common as the platy clays, do occur in many different kinds of depositional environments, including marine and lacustrine sediment, hydrothermal deposits, and soils of arid regions (Singer and Galan, 1984). So-called amorphous clay, referred to as allophane, is an aluminosilicate of variable chemical composition that is amorphous to x-rays in most cases but does diffract electrons. It may be an intermediate product in the formation of phyllosilicate clays from various kinds of materials, but its abundance is not well documented because it is not directly detectable by x-ray diffraction (Carroll, 1970).

Clay minerals are described in many textbooks on mineralogy and in numerous specialized books, including those by Grim (1968), Millot (1970), Weaver and Pollard (1973), Velde (1977), Sudo and Shimoda (1978), Sudo et al. (1981), and Cairns-Smith and Hartman (1987). In addition, several scientific journals are devoted to clay min-

eralogy, including *Clays and Clay Minerals*, *Clay Minerals Bulletin*, *Clay Science*, and *Clays and Clay Technology*. Articles on clay mineralogy are also published in the *American Mineralogist*, the *Journal of the Soil Science Society of America*, and the *Journal of Sedimentary Petrology*.

The phyllosilicate clay minerals consist of two kinds of layers of different chemical composition and coordination. One of these is the *tetrahedral layer* (T), which is composed of silica tetrahedra. The second layer consists of Al hydroxide with Al^{3+} ions located in the centers of octahedra and is therefore known as the *octahedral layer* (O). In some clay minerals, such as serpentine and talc, the octahedral layer is composed of Mg hydroxide, which forms the mineral *brucite*. Therefore, the octahedral layer is also referred to as the *brucite layer* or the *gibbsite layer*, depending on its composition.

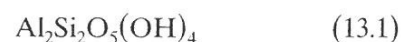
When Al^{3+} occurs in the octahedral layer, only two of every three available sites are filled. Therefore, such clays are said to be dioctahedral. When Al^{3+} is replaced by Fe^{3+} or by other trivalent ions, the minerals remain dioctahedral. However, when the two Al^{3+} ions are replaced by three divalent ions of Fe^{2+} , Mg^{2+} , Zn^{2+} , etc., then all positions in the octahedral layer are filled and the clay is trioctahedral. Based on these distinctions we conclude that kaolinite is dioctahedral, whereas serpentine is trioctahedral.

The tetrahedral layer consists of silica tetrahedra that are linked together by sharing O atoms at three of the four corners. The linkages occur only at the bases of the tetrahedra, thus forming a sheet with the unshared O atoms all pointing in the same direction. If *all* four O atoms of silica tetrahedra are shared (as in feldspar), then each Si atom controls only one half of its O atoms. Therefore, the Si/O ratio in feldspar and other tectosilicates is 1 : 2 (Section 8.3). Since in phyllosilicates only the three *basal* O atoms are shared, the Si/O ratio is 1 : 1.5. When the *unshared* O atom is counted with the basal ones, the Si/O ratio becomes 1 : 2.5.

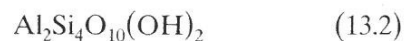
The octahedral layer has six hydroxyl ions located at the corners of an octahedron. Therefore, the ratio of Al to hydroxyl ions is basically 1 : 6. However, the octahedra that make up the gibbsite sheet are linked at the corners by

sharing hydroxyl ions and have an Al/OH ratio of 1 : 3 instead of 1 : 6.

The tetrahedral and octahedral layers are joined to form two-layer clays (T-O), three-layer clays (T-O-T), or mixed-layer clays that are mixtures of different two-layer and three-layer clays. The linking of the octahedral and tetrahedral sheets in T-O clay minerals takes place by sharing of O atoms and hydroxyl radicals, which reduces the number of hydroxyls by *one each time such a link is established*. Therefore, two-layer clays, such as kaolinite, are composed of Si, Al, O, and OH in the proportions $\text{SiO}_{2.5}\text{Al}(\text{OH})_2$. We multiply the formula by two in order to clear the decimal fraction and obtain the formula of kaolinite and its polymorphs:



Similarly, the formula of three-layer clays (T-O-T), such as pyrophyllite, is $2(\text{SiO}_{2.5})\text{Al}(\text{OH})$ or $\text{Si}_2\text{O}_5\text{Al}(\text{OH})$. Again, we multiply by two and obtain:



The structures of two- and three-layer clays are illustrated in Figure 13.1.

The phyllosilicate clay minerals have widely varying chemical compositions because the Al^{3+} ion in the octahedral layers can be wholly or partly replaced by Fe^{3+} , Cr^{3+} , Fe^{2+} , Mg^{2+} , Zn^{2+} , Li^+ , and many other ions. In addition, some of the Si^{4+} ions in the tetrahedral layer are replaced by Al^{3+} . This results in an excess of negative charge, which is neutralized by the adsorption of cations to the outer surfaces of the tetrahedral layers of adjoining clay units. The substitutions within the layers are constrained by the requirement for electrical neutrality, which is met when the cations in the *octahedral layer* have a combined charge of +6 and the *tetrahedral layers* contain four moles of Si + Al per formula weight. The two-layer clays permit very little substitution of either Al or Si, whereas the three-layer clays, with the exception of pyrophyllite, are characterized by extensive substitutions in either the octahedral layer or the tetrahedral layer, or in both.

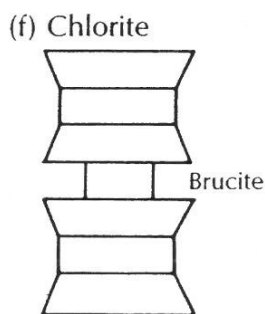
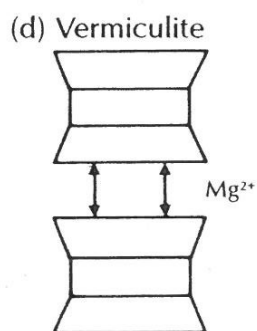
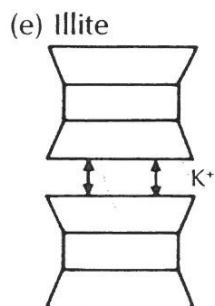
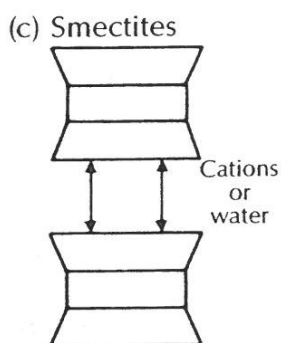
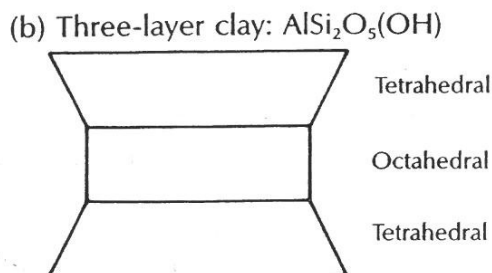
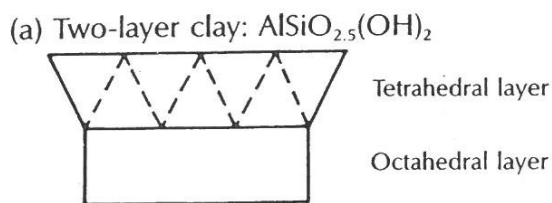


Figure 13.1 Schematic diagrams of the structures of two-layer and three-layer clay minerals.

13.2 Classification and Chemical Composition

The classification of clay minerals is based on their crystal structure and chemical composition. In practice, clay minerals are identified by x-ray diffraction combined with treatment procedures that together serve to identify certain mineral species. In other words, some clay minerals are defined by the procedure that is used to identify them.

a. Two-Layer Clays

The phyllosilicate clays are subdivided into the two-layer, three-layer, and mixed-layer clays. The two-layer clays consist of two subgroups, each of which contains a number of mineral species, as listed in Table 13.1.

1. KAOLINITE The minerals in the *kaolinite subgroup* all have the familiar formula $\text{Al}_2\text{Si}_2\text{O}_5(\text{OH})_4$ and include *kaolinite*, *dickite*, *nacrite*, and *halloysite*. These minerals differ from kaolinite only by having different crystal structures. For example, halloysite forms tubular crystals, whereas kaolinite crystals are platy, but all members of the kaolinite subgroup have identical chemical compositions. Therefore, halloysite and the other mineral species in the kaolinite subgroup are *polymorphs* of kaolinite but occur in certain geochemical environments in preference to kaolinite.

2. SERPENTINE AND GREENALITE *Serpentine* is a two-layer phyllosilicate like kaolinite in which the gibbsite layer is replaced by brucite. Consequently, the formula of serpentine $[\text{Mg}_3\text{Si}_2\text{O}_5(\text{OH})_4]$ contains Si and O in the characteristic proportion of the tetrahedral layer, and the octahedral layer contains a positive charge of +6 as a result of the replacement of two Al^{3+} by three Mg^{2+} . The mineral species in that group are *chrysotile*, *lizardite*, and *antigorite*, all of which are polymorphs of serpentine. *Greenalite* $[\text{Fe}_3^{2+}\text{Si}_2\text{O}_5(\text{OH})_4]$ is a two-layer clay in which three Fe^{2+} ions take the place of two Al^{3+} ions.

Table 13.1 Classification of Clay Minerals

Group	Subgroup	Species
A. Two-Layer Clays		
Kaolinite	kaolinite (dioctahedral)	kaolinite
		dickite
		nacrite
		halloysite
Kaolinite	serpentine (trioctahedral)	chrysotile
		lizardite
		antigorite
		greenalite
B. Three-Layer Clays		
Pyrophyllite	dioctahedral	pyrophyllite
	trioctahedral	talc
Smectite	dioctahedral	minnesotaite
	dioctahedral	montmorillonite
		beidellite
Smectite	trioctahedral	nontronite
		saponite
		hectorite
Vermiculite	dioctahedral	sauconite
	trioctahedral	vermiculite
		vermiculite
Mica	dioctahedral	illite
		muscovite
		paragonite
		glauconite
Mica	trioctahedral	biotite
		phlogopite
		lepidomelane
Brittle mica	dioctahedral	margarite
	trioctahedral	seybertite
		xanthophyllite
		Chlorite
trioctahedral	cookeite (?)	
	Fe-rich varieties	
	thuringite	
	chamosite	
	Mg-rich varieties	
	clinochlore	
	penninite	

Table 13.1 (continued)

Group	Subgroup	Species
C. Mixed-Layer Clays		
Complex interstratified two-layer and three-layer clay minerals following either a regular or a random pattern.		
D. Fibrous Clay Minerals		
		palygorskite (attapulgite)
		sepiolite

Adapted from Grim (1968) and Carroll (1970).

b. Three-Layer Clays

The three-layer clays can be divided into the *pyrophyllite*, *smectite*, *vermiculite*, *mica*, *brittle mica*, and *chlorite groups*. All of the minerals included here have the basic three-layer structure of pyrophyllite but differ from it in chemical composition and physical properties.

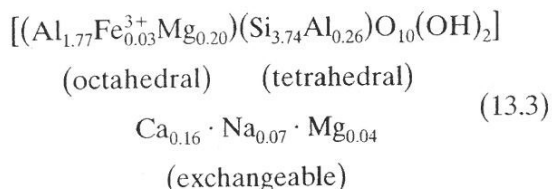
1. PYROPHYLLITE AND TALC *Pyrophyllite* itself ($\text{Al}_2\text{Si}_4\text{O}_{10}(\text{OH})_2$) is the prototype of all three-layer clays and permits no substitutions in the octahedral or tetrahedral layers.

Talc is related to pyrophyllite in the same way that serpentine is related to kaolinite because in talc the octahedral layer is composed of brucite instead of gibbsite. Consequently, talc has the formula $\text{Mg}_3\text{Si}_4\text{O}_{10}(\text{OH})_2$, where three Mg^{2+} ions replace two Al^{3+} and the Si/O ratio is derived from the presence of two tetrahedral layers. *Minnesotaite* is the Fe member in this group in which the octahedral layer contains Fe^{2+} instead of Al^{3+} . Consequently, its formula is $\text{Fe}_3^{2+}\text{Si}_4\text{O}_{10}(\text{OH})_2$.

2. SMECTITES The *smectite group* contains a large number of mineral species that differ in chemical composition because of the complete or partial replacement of Al^{3+} in the octahedral layer and the partial replacement of Si^{4+} in the tetrahedral layer.

The *smectite group* is divided into dioctahedral and trioctahedral subgroups each of which contains

mineral species that differ from each other in chemical composition. For example, *montmorillonite* is a prominent member of the *dioctahedral subgroup* of smectites in which Mg^{2+} occurs in the octahedral layer, but the replacement of Si^{4+} by Al^{3+} in the tetrahedral layer is quite limited. For example, a sample of montmorillonite from Germany has the formula (Carroll, 1970, Table 6):



This formula can be readily interpreted in terms of the three-layer model if we determine the electrical charges in each of the layers. The octahedral layer contains $1.77 \text{ Al}^{3+} + 0.03 \text{ Fe}^{3+} + 0.20 \text{ Mg}^{2+}$, which adds up to $+5.8$ electronic charges. Since the octahedral layer is supposed to contain six positive charges, there is a deficiency of 0.2 positive charge, which is equivalent to an excess of 0.2 negative charge. In the tetrahedral layer replacement of some Si^{4+} by Al^{3+} causes a deficiency of positive charge equivalent to a charge of -0.26 . Evidently, the substitutions in the two kinds of layers have caused a total charge of $-0.2 + (-0.26) = -0.46$. In the example before us, the negative charges are neutralized by adsorption of $0.16 \text{ Ca}^{2+} + 0.07 \text{ Na}^{+} + 0.04 \text{ Mg}^{2+}$, which together provide 0.47 positive charge. The agreement between the negative charges in the layers (-0.46) and the positive charges of the additional ions ($+0.47$) is well within the uncertainty of the chemical analysis of this mineral. The overall charge distribution for this montmorillonite is given by adding up the ionic charges of all elements that make up this mineral:

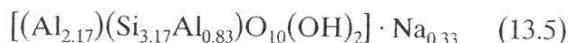
$$\begin{aligned} &1.77(\text{Al}^{3+}) + 0.03(\text{Fe}^{3+}) + 0.20(\text{Mg}^{2+}) \\ &+ 3.74(\text{Si}^{4+}) + 0.26(\text{Al}^{3+}) \\ &+ 0.16(\text{Ca}^{2+}) + 0.07(\text{Na}^{+}) \\ &+ 0.04(\text{Mg}^{2+}) - 10(\text{O}^{2-}) - 2(\text{OH}^{-}) \\ &= +0.01 \end{aligned} \quad (13.4)$$

The additional cations, which neutralize the negative charges caused by imperfect substitutions within the layers, are *adsorbed* on the surfaces of the mineral grains and on the tetrahedral layers of the clay units. Therefore, they are partly or completely *exchangeable* and can be replaced by other cations or by the polar molecules of water in the environment. The adsorbed ions located on the outer surfaces of the tetrahedral layers bond the structural units together. When the bonding involves the cations of metals, the interlayer distances between adjacent structural units are short. However, when certain clay minerals are immersed in water, the interlayer cations are replaced by molecules of water, which have much smaller net positive charges than metal ions. Consequently, a larger number of water molecules is required to neutralize excess negative charges generated within the layers. As a result, the interlayer distances *increase* as water molecules crowd in and the clay *expands*.

The swelling of clay minerals in the smectite group is a characteristic property that is used to identify them. For example, when natural montmorillonite is heated at 300°C for an hour or more, the interlayer distance *decreases* from about 15 angstroms ($\text{\AA}$) to about 9 $\text{\AA}$ because adsorbed water molecules are expelled. When the clay is subsequently exposed to the vapor of ethylene glycol ($\text{HOCH}_2\text{CH}_2\text{OH}$) for several hours, the interlayer distance *expands* to 17 or 18 $\text{\AA}$.

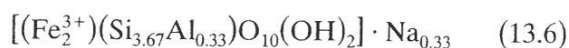
Montmorillonite commonly forms by alteration of volcanic ash and therefore is the principal mineral of *bentonite* beds, which form by alteration of volcanic ash deposits. The well-known swelling of bentonite, when it is placed in water, is actually caused by the expansion of the interlayer distances of montmorillonite. The mineral also occurs in soils of arid regions where leaching is limited by lack of water and the pH is above 7 .

Beidellite is a dioctahedral smectite containing little if any Fe^{2+} or Mg^{2+} and in which negative charges originate primarily by replacement of Si^{4+} by Al^{3+} in the tetrahedral layer. According to Grim (1968), its formula can be written as:



Note that the replacement of Si^{4+} by Al^{3+} in the tetrahedral layer has caused a net charge of -0.83 , which is partly compensated by the positive charge ($+0.51$) generated by an excess of 0.17 Al^{3+} in the octahedral layer. The remaining excess negative charge ($-0.83 + 0.51 = -0.32$) is neutralized by adsorption of 0.33 moles of Na^+ per formula weight of beidellite.

Nontronite, the third dioctahedral smectite in Table 13.1, contains Fe^{3+} in the octahedral layer in place of Al^{3+} with minor amount of Fe^{2+} and Mg^{2+} . The net charge imbalance arises from substitutions of Si^{4+} by Al^{3+} in the tetrahedral layer. Therefore, the formula of nontronite is:

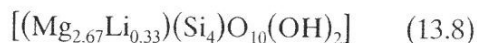


As in beidellite, the negative charge generated in the tetrahedral layer may be partly compensated by the presence of additional Fe^{3+} , Fe^{2+} , or Mg^{2+} in the octahedral layer. When Al^{3+} in the octahedral layer is replaced by Cr^{3+} , the resulting dioctahedral smectite is known as *volkhonskoite*.

The *trioctahedral smectites* include the mineral species *saponite*, *hectorite*, and *sauconite*. In *saponite*, the octahedral layer is composed primarily of Mg^{2+} in place of Al^{3+} with minor amounts of Fe^{3+} and Fe^{2+} . Negative charges arise in the tetrahedral layer by replacement of Si^{4+} by Al^{3+} and may be partly compensated by additional cations in the octahedral layer. The formula of *saponite* according to Grim (1968) is:

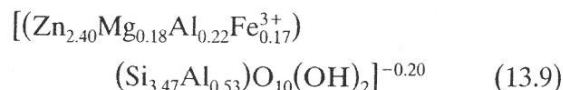


Hectorite is a Li-bearing smectite in which Mg^{2+} and Li^+ occur together in the octahedral layer, but no substitutions occur in the tetrahedral layer. Thus the formula for *hectorite* is:



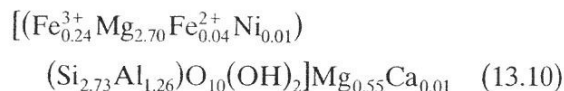
The smectite mineral *stevensite* is similar to *hectorite* because it is Mg-rich and lacks Al in the tetrahedral layer. However, *stevensite* also lacks Li in the octahedral layer.

The smectite mineral *sauconite* is characterized by the presence of Zn^{2+} in the octahedral layer and by limited replacement of Si^{4+} by Al^{3+} in the tetrahedral layer. Carroll (1970) recalculated a chemical analysis of *sauconite* to obtain the following formula:



The charge imbalance in the octahedral layer of this clay mineral is $+0.33$, that of the tetrahedral layer is -0.53 for a net charge of -0.20 , which is neutralized by the adsorption of exchangeable cations in interlayer sites.

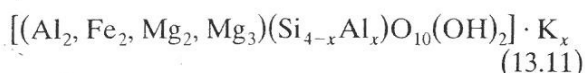
The word *vermiculite* is derived from the Latin verb *vermiculari*, which means to *breed worms*. The clay mineral *vermiculite* occurs both in dioctahedral and trioctahedral varieties with other clay minerals and as an alteration product of micas. According to Gruner (1934), *vermiculite* from Bare Hills, Maryland, has the formula:



Evidently, Mg^{2+} dominates in the octahedral layer, and the Si^{4+} in the tetrahedral layer is extensively replaced by Al^{3+} . The octahedral layer has an excess positive charge of $+0.22$, which partly compensates for the strong negative charge of -1.30 in the tetrahedral layer. The remaining negative charge of -1.08 is neutralized by adsorbed Mg^{2+} and Ca^{2+} in the interlayer position. *Vermiculite* characteristically contains both Mg^{2+} and water molecules in the interlayer sites. The water can be expelled by heating at 500°C , which causes the interlayer distance to decrease from about $14 \text{ \AA}$ to about $9 \text{ \AA}$. After heating at 500°C or less, *vermiculite* rehydrates spontaneously; however, heating at 700°C causes irreversible dehydration. In addition, when *vermiculite* is treated with a solution of KCl, the interlayer Mg^{2+} ions are replaced by K^+ , and it becomes nonexpandable. *Vermiculite* differs from smectites by having a higher charge in the tetrahedral layer, by commonly containing

Mg²⁺ in the exchangeable interlayer sites, and by being less expandable.

3. ILLITE The power of K⁺ ions to stabilize clay minerals is well illustrated by *illite*, which is listed in Table 13.1 with the *mica* minerals. Actually, illite could also be regarded as a special member of the *smectite* group. However, it differs from typical smectites by having a *nonexpandable lattice* because of the presence of K⁺ ions in interlayer sites. Illite was named by Grim et al. (1937) after the state of Illinois in the United States. They intended it to represent mica-like clay minerals as a group, rather than a specific clay mineral, and expressed the formula of illite as:



The formula indicates that the octahedral layer may contain Al, Fe, or Mg and that illites may be dioctahedral (2 Al³⁺ or 2 Fe³⁺) or trioctahedral (3 Mg²⁺). The value of x is less than one and usually varies between 0.5 and 0.75. The negative charge generated in the tetrahedral layer is compensated by K⁺ ions, which fit into hexagonal holes formed by the silica tetrahedra of the tetrahedral layers. As a result, the K⁺ ions are not exchangeable, the spacing of the interlayers is fixed, and the mineral is not expandable. However, illites may be interlayered with montmorillonite or other clay minerals that are expandable.

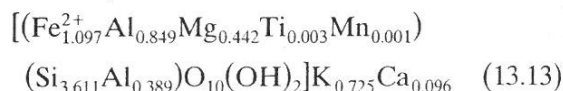
Illite is a common weathering product of *muscovite* and differs from it primarily by the extent of substitution of Si⁴⁺ by Al³⁺ in the tetrahedral layer. *Muscovite* and its Na-analog *paragonite* are dioctahedral and have the formula:



The charge deficiency in the tetrahedral layer of -1 is neutralized by K⁺ in *muscovite* and by Na⁺ in *paragonite*. The stacking of adjacent mica units in *muscovite* can occur in different ways, giving rise to polymorphs designated as 1M, 1Md, 2M₁, 2M₂, and 3T (Smith and Yoder, 1956). Illite in the smallest grain-size fraction (<0.2 μm) of sediment or soil is of the 1Md or 1M type. Larger grains consist of the

2M polymorphs and are derived by mechanical weathering of igneous and metamorphic rocks. Therefore, the 2M illite in sedimentary rocks is *detrital* in origin, whereas the 1Md or 1M polymorphs may consist of *authigenic* illite. The x-ray diffraction spectra of the different muscovite polymorphs in the clay fraction of sedimentary materials were worked out by Yoder and Eugster (1955).

4. GLAUCONITE OR GLAUCONY *Glaucinite* is a dioctahedral clay mineral in which Fe³⁺, Fe²⁺, and Mg²⁺ replace Al³⁺ in the octahedral layer. Negative charges in the octahedral and tetrahedral layers are neutralized by interlayer K⁺, Ca²⁺, and Na⁺. The concentration of K⁺ in glauconite is variable and increases with time depending on the availability of K⁺ in the environment. Johnston and Cardile (1987) reported that glauconite from Point Jackson, Francosia, Wisconsin, has the formula:



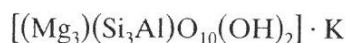
In this sample of glauconite, Fe²⁺ dominates in the octahedral layer, and the deficiency of positive charges in both the octahedral and tetrahedral layers is largely compensated by interlayer K⁺ and Ca²⁺ ions.

Glaucinite is an *authigenic* clay mineral that forms in the marine environment by transformation from other kinds of material (McRae, 1972). It commonly occurs as greenish pellets about 1–2 mm in diameter comprised of aggregates of platy glauconite crystallites about 10 μm in diameter. Glaucinite has been used for dating unmetamorphosed sedimentary rocks by the K–Ar and Rb–Sr methods (Chapter 16) because it contains both K and Rb and because it is *authigenic* in origin.

The status of glauconite as a discrete mineral species has become controversial, partly because of its importance for dating sedimentary rocks and the hoped-for refinement of the geological time scale. Burst (1958) and Hower (1961) regarded glauconite as an Fe-rich illite with variable K content that occurs in pure form as the 1Md or 1M polymorph. However, this definition included poorly crystallized materials with low K concentrations that are

unsuitable for dating by isotopic methods. Therefore, Odin and Matter (1981) proposed the word "glaucy" for a spectrum of Fe-rich clay minerals with varying K concentrations that form the characteristic green pellets in sedimentary rocks of marine origin. Glaucy therefore represents clay minerals that are intermediate between Fe-rich and K-poor smectite and Fe-rich K-mica. As the K concentration of glaucy increases, the material becomes less expandable and more retentive with respect to the radiogenic daughters of K and Rb.

5. MICAS The trioctahedral micas in igneous and metamorphic rocks include *biotite* and *phlogopite*. These differ from muscovite by the chemical composition of the octahedral layer, but both replace one Si^{4+} by one Al^{3+} in the tetrahedral layer. Actually, biotite is intermediate between phlogopite:



and annite or lepidomelane:



Therefore, the general formula for biotite can be written as:



Biotite is considerably more susceptible to chemical weathering than muscovite because the Fe^{2+} can oxidize to Fe^{3+} in the presence of an electron acceptor. As a result, the charge balance of the lattice is disturbed and the mineral alters to other aluminosilicates and ferric oxide, depending on environmental conditions.

The so-called *brittle micas* in Table 13.1 include the dioctahedral species *margarite* and other rare trioctahedral phyllosilicates listed by Grim (1968).

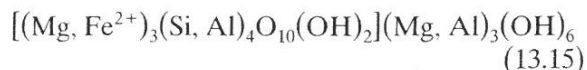
6. CHLORITE The *chlorite* group contains primarily trioctahedral species composed of three-layer clay units bonded to brucite sheets. The chemical compositions of chlorite range from Mg-rich to Fe-rich varieties, but all contain varying amounts of Al, which replaces Si in the tetrahedral layers.

The only dioctahedral chlorite is an Al-rich variety known as *cookeite* described by Brown

and Bailey (1962) and by Eggleston and Bailey (1967). Dioctahedral chlorite has been identified in soil samples with $\text{pH} < 5.5$ and may also crystallize metastably in sedimentary environments as an authigenic mineral (Carroll, 1970).

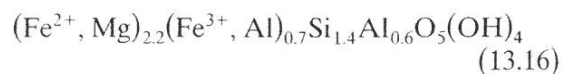
The trioctahedral chlorites have been classified on the basis of their chemical compositions by Hey (1954) and by Foster (1962). The latter specified compositional limits for eight varieties of chlorite including the Fe-rich species *thuringite* and *chamosite* and the Mg-rich species *clinocllore* and *penninite*.

The structural formula of trioctahedral chlorites includes the mica unit $(\text{Mg}, \text{Fe}^{2+})_3(\text{Si}, \text{Al})_4\text{O}_{10}(\text{OH})_2$, which carries a negative charge because of the replacement of up to 1.6 Si^{4+} by Al^{3+} in the tetrahedral layer. The negative charge is neutralized by a brucite sheet that carries a positive charge caused by replacement of Mg^{2+} by Al^{3+} : $(\text{Mg}, \text{Al})_3(\text{OH})_6$. The brucite layer occupies an interlayer position and bonds two mica units together as K^+ ions do in illite (Figure 13.1). The complete structural formula of chlorite is obtained by combining the components of the mica and brucite units:



In addition, chlorites may contain Fe^{3+} ions, which replace Al^{3+} in the brucite layer.

Chamosite, one of the Fe-rich chlorites, is an important constituent of Fe-rich sedimentary rocks of marine origin described by James (1966). Another Fe-rich phyllosilicate called *berthierine* has a 1 : 1 kaolinite-type lattice, but has also been called "chamosite" (Brindley, 1961, Brindley et al., 1968). This mineral has the formula:



and is structurally related to kaolinite but compositionally similar to serpentine. However, it differs from serpentine by containing Fe in the octahedral layer and by limited substitution of Si^{4+} by Al^{3+} in the tetrahedral layer.

Both chamosite and berthierine are authigenic minerals that form in the ocean and may

recrystallize during low-grade regional metamorphism to the Fe-chlorite *thuringite*.

c. Mixed-Layer Clays

The clay minerals we have described occur as interstratified layers in the so-called mixed-layer clays, which are very common in soils and in about 70% of clay-rich sedimentary rocks (Weaver, 1956). To some extent the principal types of mixed-layer clays depend on climatic factors that affect the environment within which chemical weathering occurs. Under moist and tropical conditions the stratification in mixed-layer clays is montmorillonite-halloysite-kaolinite, whereas in humid temperature regions the mixed-layer clays consist of montmorillonite-chlorite-mica or mica-intermediate products-illite (Carroll, 1970). The crystallographic properties of mixed-layer clays have been discussed by Zen (1967), Reynolds (1967, 1980), and Corbató and Tettendorst (1987).

d. Fibrous Clay Minerals

The fibrous clay minerals palygorskite and sepiolite are composed of double chains of silica tetrahedra with an Si/O ratio of 4:11. The channel between the double chains is occupied by water molecules that are lost stepwise during heating up to 850°C, at which temperature the structure is destroyed. Neither palygorskite nor sepiolite expand when treated with organic liquids.

Palygorskite is a Mg-silicate with some replacement of Mg^{2+} and Si^{4+} by Al^{3+} . Its structure was worked out by Bradley (1940), who reported the formula:

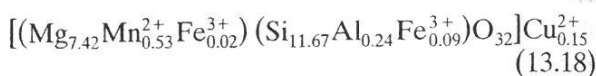


where OH_2 is bound water and H_2O is water in channels like that of zeolites. The structure presented above is electrically neutral but may contain exchangeable Ca^{2+} to compensate charge imbalances caused by limited substitutions in this structure. Palygorskite is called *attapulgit* in the United States after an occurrence of this clay in Attapulgis, Georgia.

Palygorskite is an authigenic mineral that occurs in alkaline, unleached soils and as a product

of hydrothermal alteration of pyroxene and amphibole. One of the largest accumulations is in the Hawthorn Formation of Miocene age in Georgia and Florida, which was deposited in fresh to brackish water lagoons.

Sepiolite is also composed of chains of silica tetrahedra but it differs from palygorskite in specific details. Grim (1968) reported two different structural formulas for this mineral based on a review of the literature available at the time. The formulas currently in use (see Appendix B) are $Mg_2Si_3O_6(OH)_4$ and $Mg_4Si_6O_{15}(OH)_2 \cdot 6H_2O$. Carroll (1970) reported that a chemical analysis of sepiolite from Little Cottonwood, Utah, yields the formula:



Evidently, Mg^{2+} is partly replaced by Mn^{2+} and Fe^{3+} , whereas Si^{4+} is replaced by Al^{3+} and Fe^{3+} . The structure has a net charge of -0.37 , which is largely neutralized by exchangeable Cu^{2+} .

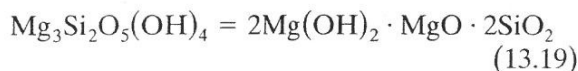
13.3 Gibbs Free Energies of Formation

The thermodynamic properties of the three-layer and mixed-layer clay minerals are difficult to establish because of their highly variable chemical compositions. As a result, standard free energies or enthalpies of formation are available only in certain specific cases or apply only to idealized compositions that do not reflect the true chemical diversity of real clay minerals.

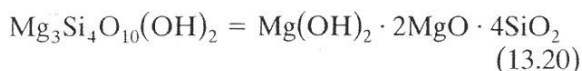
In order to overcome this problem, Tardy and Garrels (1974), Chen (1975), and Nriagu (1975) proposed empirical methods for estimating standard free energies of formation of clay minerals. Tardy and Garrels (1976, 1977), Tardy and Gartner (1977), Tardy and Vieillard (1977), and Tardy (1979) subsequently generalized this approach to estimate the standard free energies of formation of other kinds of compounds and their ions in aqueous solution.

The method of Tardy and Garrels (1974) is based on the premise that the standard free energy of a phyllosilicate can be treated as the sum of

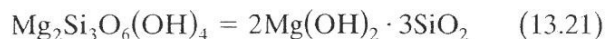
the free energies of its oxide and hydroxide components. The numerical values of the free energies of the components must reflect the fact that they exist in a silicate lattice and not in free form. Therefore, they must first be determined from phyllosilicate minerals whose standard free energies of formation are known from experimental determinations. The first step toward that end is to restate the structural formulas of phyllosilicates in terms of oxide components, except that Mg is converted to the hydroxide until all of the hydroxyl ion in the formula has been assigned to $\text{Mg}(\text{OH})_2$. For example, serpentine is rewritten:



Similarly, talc becomes:



and sepiolite, which is not a phyllosilicate, nevertheless is recast into the form:



The three Mg-silicates are composed of $\text{Mg}(\text{OH})_2$, MgO , and SiO_2 , whose standard free energies in the silicate lattice are to be determined. This is done by assuming that the known standard free energies of the three minerals can be expressed as the sums of the unknown standard free energies of the components in the silicate lattice. Therefore, we can write appropriate equations for each of the three minerals. For serpentine:

$$\begin{aligned} 2G_f^\circ(\text{Mg}(\text{OH})_2^s) + G_f^\circ(\text{MgO}^s) + 2G_f^\circ(\text{SiO}_2^s) \\ = -965.1 \text{ kcal/mol} \end{aligned} \quad (13.22)$$

For talc:

$$\begin{aligned} G_f^\circ(\text{Mg}(\text{OH})_2^s) + 2G_f^\circ(\text{MgO}^s) + 4G_f^\circ(\text{SiO}_2^s) \\ = -1320.38 \text{ kcal/mol} \end{aligned} \quad (13.23)$$

And for sepiolite:

$$\begin{aligned} 2G_f^\circ(\text{Mg}(\text{OH})_2^s) + 3G_f^\circ(\text{SiO}_2^s) \\ = -1020.95 \text{ kcal/mol} \end{aligned} \quad (13.24)$$

Equations 13.22–13.24 can be solved by substitution and yield the standard free energies of formation in the silicate lattice of the three components listed in Table 13.2. Other components can be added by considering the Al silicates kaolinite, pyrophyllite, and muscovite, which yield values for $G_f^\circ(\text{Al}_2\text{O}_3^s)$, $G_f^\circ(\text{H}_2\text{O}^s)$, and $G_f^\circ(\text{K}_2\text{O}^s)$. Tardy and Garrels (1974) also determined the standard free energies of formation in the silicate lattice of exchangeable cations and water. These values are listed in Table 13.2.

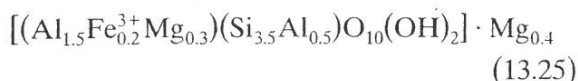
In order to illustrate their method of estimating standard free energies of formation of a real

Table 13.2 Standard Free Energies of Formation of Oxide and Hydroxide Components in the Lattice of Phyllosilicate Minerals at 25°C

Component	G_f° (silicated), kcal/mol
K_2O	-188.0
Na_2O	-162.8
MgO	-149.2
FeO	-64.1
Fe_2O_3	-177.7
Al_2O_3	-382.4
SiO_2	-204.6
H_2O	-59.2
$\text{Mg}(\text{OH})_2$	-203.3
$\text{Al}(\text{OH})_3$	-280.0
FeOOH	-117.0
KOH	-123.6
NaOH	-111.0
<i>Exchangeable Ions Expressed as Oxides</i>	
K_2O	-188.0
CaO	-182.8
Na_2O	-175.4
MgO	-159.5
H_2O	-58.6
Li_2	-190.6

SOURCE: Tardy and Garrels (1974).

clay mineral, Tardy and Garrels (1974) chose a typical montmorillonite:



The formula is electrically neutral, which is important because the exchangeable cations do contribute to the free energy of the mineral and therefore cannot be omitted. In Table 13.3 the formula for this montmorillonite has been recast in terms of molar amounts of the oxide components. These are then multiplied by their respective G_f° (silicated) values from Table 13.2 and the products are summed. The result is -1282.5 kcal/mol, which is a reasonable value for the standard free energy of formation of montmorillonite. For example, Appendix B contains a value of -1258.84 kcal/mol listed by Lindsay (1979) for a Mg-montmorillonite.

More specific comparisons between free energies of formation, estimated by the method of Tardy and Garrels (1974), and experimentally determined values are presented in Table 13.4. The experimental determinations were made by

Table 13.3 Estimate of the Standard Free Energy of Formation of a Typical Montmorillonite with the Formula Stated in Equation 13.25

1 Component, moles	2 G_f° (silicated), kcal/mol	3 1×2 , kcal
1.0 Al_2O_3	-382.4	-382.4
0.1 Fe_2O_3	-177.7	-17.77
0.3 $\text{Mg}(\text{OH})_2$	-203.3	-60.99
3.5 SiO_2	-204.6	-716.1
0.7 H_2O^a	-59.2	-41.44
0.4 $\text{MgO}(\text{exch.})$	-159.5	-63.8
	Sum	-1282.5 kcal

^a0.3 $\text{Mg}(\text{OH})_2$ uses up 0.6 hydrogen atom, leaving 1.4 hydrogen atoms to make 0.7 mole of H_2O .

SOURCE: Tardy and Garrels (1974).

Table 13.4 Standard Free Energies of Formation of Clay Minerals Based on Experimental Determinations and Estimates by the Method of Tardy and Garrels (1974)

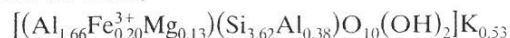
Mineral or locality	G_f° (measured), kcal/mol	G_f° (estimated), kcal/mol
Illite		
Fithian	-1277.2 ^a	-1278.8
Beavers Bend	-1274.7 ^a	-1276.2
Goose Lake	-1272.1 ^a	-1273.4
Montmorillonite		
Aberdeen	-1230.6 ^c	-1228.3
Belle Fourche	-1240.6 ^b	-1236.4
Colony, Wyoming	-1241.5 ^d	-1240.6
Other Phyllosilicates		
Annite (1)	-1149.3	-1150.2
Annite (2)	-1151.7	-1149.1
Clinochlore	-1961.8	-1958.7
Phlogopite	-1410.9	-1400.7
Greenalite	—	-720.0
Minnesotaite	—	-1055.0

Illite formulas

Fithian,



Beavers Bend,

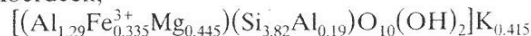


Goose Lake,

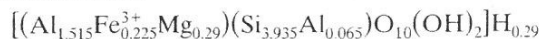


Montmorillonite formulas

Aberdeen,



Belle Fourche,



Colony,



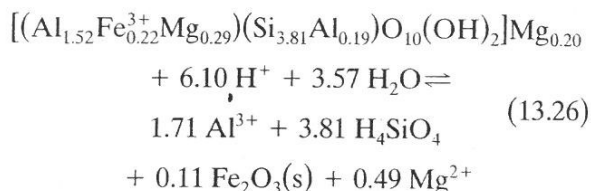
^aRoutson and Kittrick (1971).

^bKittrick (1971a).

^cKittrick (1971b).

^dWeaver et al. (1971).

dissolving the minerals and by analyzing the resulting saturated solutions. For example, Weaver et al. (1971) dissolved a sample of montmorillonite from Colony, Wyoming:



At equilibrium in a saturated solution the Law of Mass Action applies:

$$\frac{[\text{Mg}^{2+}]^{0.49}[\text{H}_4\text{SiO}_4]^{3.81}[\text{Al}^{3+}]^{1.71}}{[\text{H}^+]^{6.10}} = K \quad (13.27)$$

where $\text{Fe}_2\text{O}_3(\text{s})$ and $\text{H}_2\text{O}(\text{l})$ were given activities of one. The value of the equilibrium constant, calculated from equation 13.27 using the measured activities of the ions and molecules, was found to be $K = 10^{0.991}$. Since $\log K = -\Delta G_R^\circ/1.364$, it follows that $\Delta G_R^\circ = -1.352$ kcal. The standard free energy change that takes place during the dissociation of the montmorillonite from Colony, Wyoming, in accordance with equation 13.26 can be expressed as:

$$\begin{aligned} \Delta G_R^\circ &= 1.71 G_f^\circ(\text{Al}^{3+}) + 3.81 G_f^\circ(\text{H}_4\text{SiO}_4) \\ &+ 0.11 G_f^\circ(\text{Fe}_2\text{O}_3) + 0.49 G_f^\circ(\text{Mg}^{2+}) \\ &- 3.57 G_f^\circ(\text{H}_2\text{O}) - 6.10 G_f^\circ(\text{H}^+) - G_f^\circ(\text{mont.}) \end{aligned} \quad (13.28)$$

Since ΔG_R° for this reaction is known from the measured equilibrium constant, G_f° of the montmorillonite can be calculated from the G_f° values of the other constituents. Tardy and Garrels (1974) obtained a value of $G_f^\circ = -1241.5$ kcal per mole for the Colony montmorillonite from the solubility data, whereas their own method yields $G_f^\circ = -1240.6$ kcal per mole. The agreement is excellent. However, Weaver et al. (1971) obtained $G_f^\circ(\text{mont.}) = -1255.8 \pm 0.6$ kcal/mol for the Mg-substituted montmorillonite. The

other direct comparisons in Table 13.4 are also impressive.

13.4 Stability Diagrams

The three-layer clays are more sensitive to the geochemical environment in which they exist than most other minerals because of the presence of exchangeable ions and molecules. Their approach to equilibrium takes place by exchange of the adsorbed ions with ions or molecules from the environment, by transformations involving the chemical composition of the layers and their stacking, and perhaps even by the formation of mixed-layer clays. Although ion exchange is rapid, reactions that require adjustments in the chemical compositions of the octahedral and tetrahedral layers, and in the bonding between individual clay units, are slow. Consequently, natural materials commonly contain mixtures of several clay minerals that cannot be in equilibrium with each other under one set of environmental conditions. For this reason, we now turn to the construction and use of activity diagrams for clay minerals in order to identify the stable phases and to indicate the direction of reactions among the clay minerals in a given geochemical environment.

a. Aluminosilicate Clay Minerals

In order to simplify the construction of stability diagrams, we restrict ourselves to a suite of montmorillonites in which negative charges in the tetrahedral layers are compensated by different adsorbed cations, whereas the octahedral layers consist of unaltered gibbsite sheets. These idealized montmorillonites and their standard free energies of formation were proposed by Helgeson (1969) and Helgeson et al. (1978), and are listed in Table 13.5 together with a standardized illite. These data permit us to explore the stabilities of montmorillonites in the same coordinates we used in Chapter 12 to depict the stability fields of other aluminosilicate minerals. Real montmorillonites and other smectites should react similarly after

Table 13.5 Formulas and Standard Free Energies of Formation of Standardized Montmorillonites and Illite

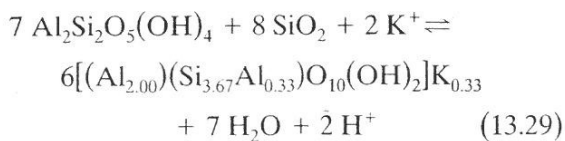
Formula	G_f° kcal/mol
Montmorillonite	
$[(Al_{2.00})(Si_{3.67}Al_{0.33})O_{10}(OH)_2]Ca_{0.167}$	-1279.24
$[(Al_{2.00})(Si_{3.67}Al_{0.33})O_{10}(OH)_2]Na_{0.33}$	-1277.76
$[(Al_{2.00})(Si_{3.67}Al_{0.33})O_{10}(OH)_2]K_{0.33}$	-1279.60
$[(Al_{2.00})(Si_{3.67}Al_{0.33})O_{10}(OH)_2]Mg_{0.167}$	-1275.34
Illite	
$[(Al_{1.80}Mg_{0.25})(Si_{3.50}Al_{0.50})O_{10}(OH)_2]K_{0.33}$	-1300.98 (-1307.485) ^a

^aEstimated by the method of Tardy and Garrels (1974).

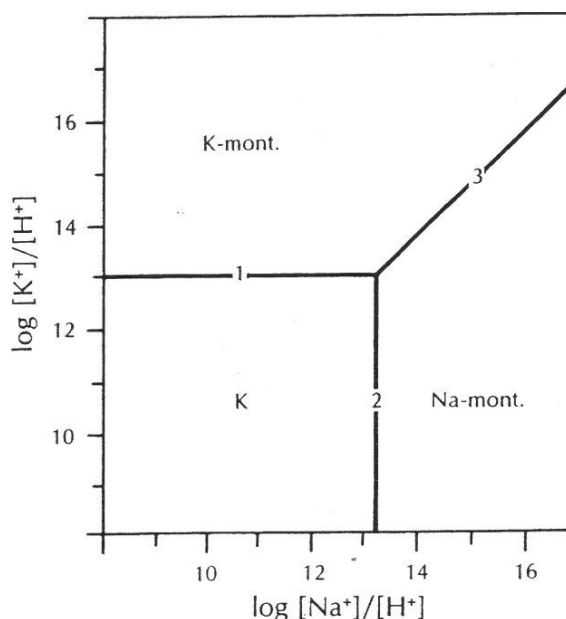
SOURCE: Helgeson (1969).

allowance is made for their more complex chemical compositions.

Based on the stability diagrams we derived in Chapter 12, we can say that kaolinite is the stable aluminosilicate clay mineral in environments where product ions are removed by leaching. If the activity of H_4SiO_4 is less than $10^{-4.7}$ mol/L, even kaolinite breaks down to form gibbsite. Therefore, we can reverse the decomposition of aluminosilicate minerals and inquire how montmorillonite and illite might form from kaolinite when the activities of their ions increase in solutions that are saturated with respect to amorphous silica. Figure 13.2 is an example of such a stability diagram in coordinates of $\log [K^+]/[H^+]$ and $\log [Na^+]/[H^+]$, assuming saturation with respect to amorphous silica. The equations for the boundaries are based on the thermodynamic data in Table 13.5 and in Appendix B. The reaction to form K-montmorillonite from kaolinite is easily constructed based on the conservation of Al:



If we multiply kaolinite by 7 and K-montmorillonite by 6, we can balance Al to a good approximation. As a result, we must add 8 moles of SiO_2 (amorphous)

**Figure 13.2** Stability of kaolinite, K-montmorillonite, and Na-montmorillonite in the presence of amorphous silica based on data in Table 13.5 and Appendix B.

to the left side of the equation to balance Si and 2 moles of K^+ to balance K. Finally, we need 7 moles of H_2O on the right side to balance O and 2 moles of H^+ to balance H. Equation 13.29 yields

$\Delta G_R^\circ = +35.511 \text{ kcal}$ and $K = 10^{-26.03}$. From the Law of Mass Action:

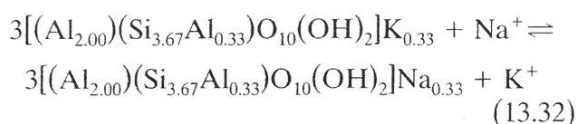
$$\frac{[\text{H}^+]^2}{[\text{K}^+]^2} = 10^{-26.03} \quad (13.30)$$

and therefore (line 1, Figure 13.2):

$$\log \frac{[\text{K}^+]}{[\text{H}^+]} = 13.02 \quad (13.31)$$

Equation 13.29 indicates that K-montmorillonite is favored over kaolinite in environments where $[\text{H}^+]$ is low and $[\text{K}^+]$ is high. Such conditions occur in poorly drained soils under arid climatic conditions or in the oceans where seawater provides a source of K^+ at a pH of about 8.2. The formation of Na-montmorillonite from kaolinite takes place by a reaction that is analogous to equation 13.29 (line 2, Figure 13.2, $\log [\text{Na}^+]/[\text{H}^+] = 13.32$).

The direct conversion of K-montmorillonite to Na-montmorillonite involves only an exchange of the two elements:



$$\Delta G_R^\circ = +0.413 \text{ kcal} \quad K = 10^{-0.302}$$

Therefore, at equilibrium:

$$\frac{[\text{K}^+]}{[\text{Na}^+]} = 10^{-0.302} \quad (13.33)$$

and hence (line 3, Figure 13.2):

$$\log \frac{[\text{K}^+]}{[\text{H}^+]} = \log \frac{[\text{Na}^+]}{[\text{H}^+]} - 0.30 \quad (13.34)$$

We can also construct a stability diagram in coordinates of $\log [\text{K}^+]/[\text{H}^+]$ and $\log [\text{Mg}^{2+}]/[\text{H}^+]^2$, which includes kaolinite, illite, muscovite, microcline, phlogopite, and chlorite in contact with amorphous silica. The equations for such a diagram are listed in Table 13.6 and have been plotted in Figure 13.3.

During the construction of Figure 13.3 we encountered difficulties caused by the lack of internal consistency of some of the thermodynamic data. A difference of 1.364 kcal in the calculated

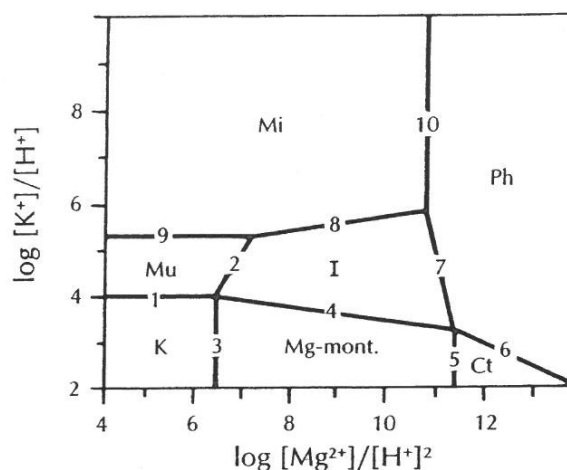


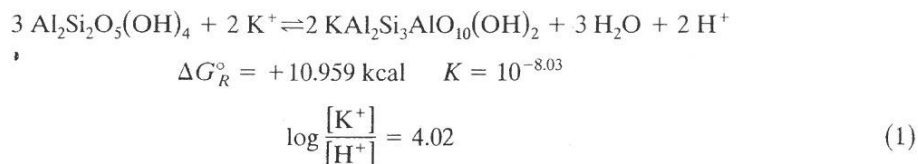
Figure 13.3 Stability of selected minerals in the system $\text{K}_2\text{O}-\text{MgO}-\text{Al}_2\text{O}_3-\text{SiO}_2-\text{H}_2\text{O}-\text{HCl}$ at 25°C at 1 atm pressure in the presence of amorphous silica. The positions of the lines were adjusted as indicated in the footnotes to Table 13.6 to overcome internal inconsistencies in the thermodynamic data. Nevertheless, the diagram shows that the clay minerals kaolinite (K), illite (I), montmorillonite (Mg-mont.), and chlorite (Ct) can form as weathering products of microcline (Mi), muscovite (Mu), and phlogopite (Ph), which crystallize from cooling magma or grow by metamorphic recrystallization of clay minerals at elevated temperatures and pressures.

value of ΔG_R° of a reaction changes its equilibrium constant by a factor of 10 and displaces the corresponding stability boundary in the activity diagram. In order to avoid some of the resulting inconsistencies, the standard free energy of formation of illite of Helgeson (1969) in Table 13.5 was recalculated by the method of Tardy and Garrels (1974), which increased it from -1300.98 to $-1307.49 \text{ kcal/mol}$. In addition, the value of G_f° of Mg-montmorillonite was changed from -1275.34 to $-1279.61 \text{ kcal/mol}$. These and other adjustments are listed at the bottom of Table 13.6.

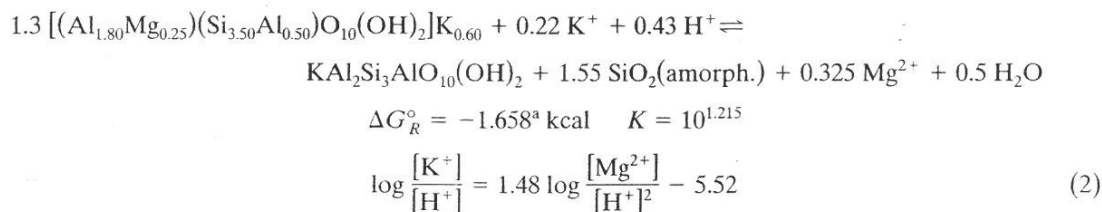
The resulting stability diagram in Figure 13.3 indicates that in acidic low-K environments phlogopite (or biotite) alters to chlorite, Mg-montmorillonite, or kaolinite, depending on the activity of Mg^{2+} . Although phlogopite may be transformed directly into each of the alteration products, sys-

Table 13.6 Equations for Mineral Stability Diagram (Figure 13.3) in the System K_2O - MgO - Al_2O_3 - SiO_2 - HCl - H_2O at 25 °C, 1 atm Pressure, and in Equilibrium with Amorphous Silica

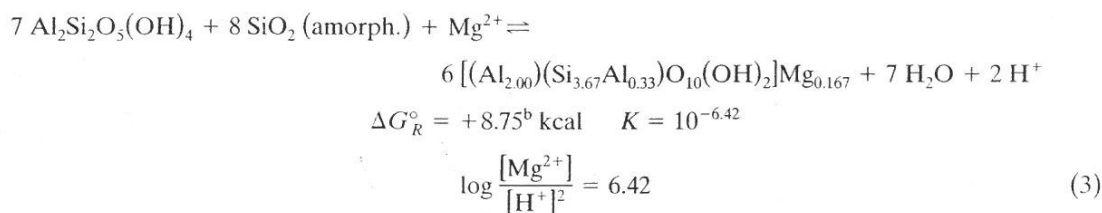
Kaolinite-muscovite



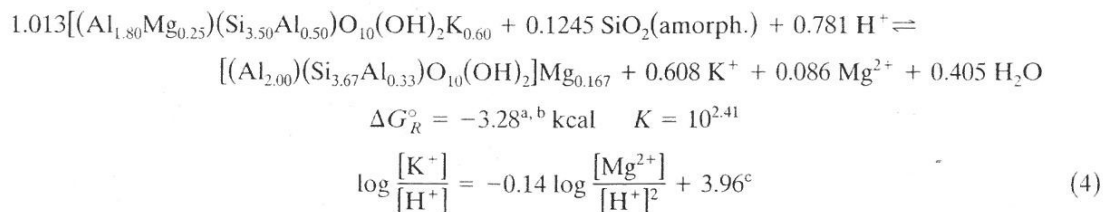
Illite-muscovite



Kaolinite-Mg-montmorillonite



Illite-Mg-montmorillonite



Mg-montmorillonite-chlorite

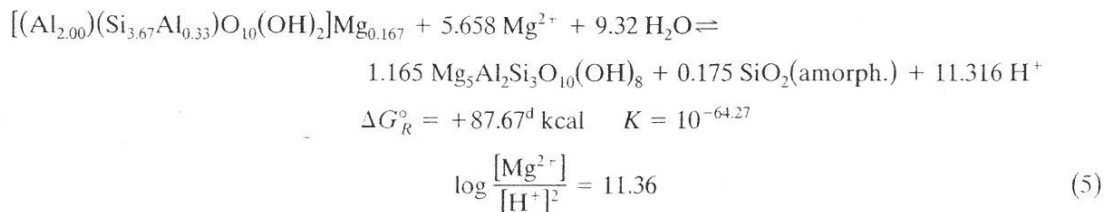
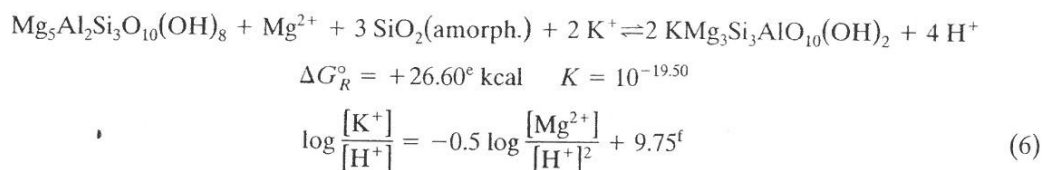
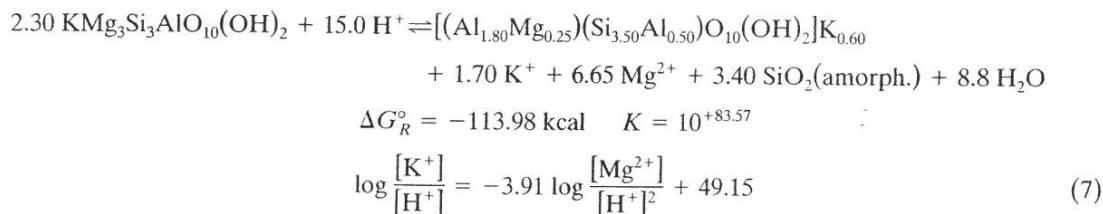


Table 13.6 (continued)

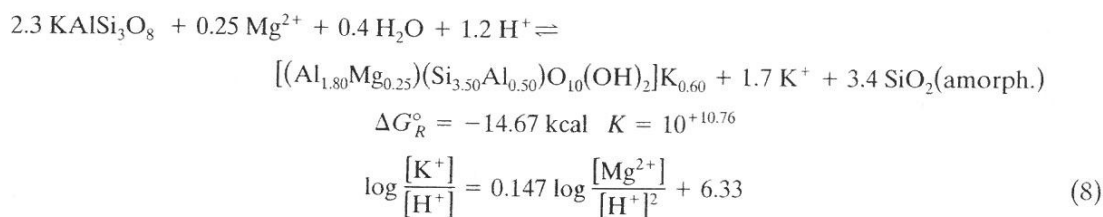
Chlorite-phlogopite



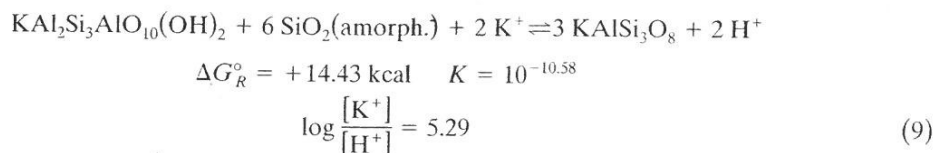
Phlogopite-illite



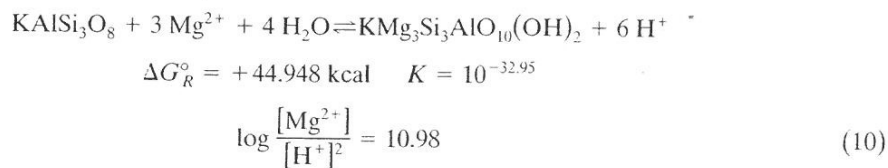
Microcline-illite



Muscovite-microcline



Microcline-phlogopite



^aBased on G_f° (illite) = -1307.485 kcal/mol, according to Tardy and Garrels (1974).

^b G_f° (Mg-montmorillonite) = -1279.61 kcal/mol adjusted so that the kaolinite-Mg-montmorillonite line goes through $\log [\text{K}^+]/[\text{H}^+] = 4.02$ and $\log [\text{Mg}^{2+}]/[\text{H}^+]^2 = 6.42$.

^cIncreased intercept to 4.86 to force line through $\log [\text{K}^+]/[\text{H}^+] = 4.02$ and $\log [\text{Mg}^{2+}]/[\text{H}^+]^2 = 6.42$.

^d G_f° (chlorite) = -1974.0 kcal/mol (Zen, 1972).

^e G_f° (phlogopite) = -1400.7 kcal/mol (Tardy and Garrels, 1974).

^fReduced to 9.0 to improve the fit.

tematic variations in the chemical composition of the fluid in the immediate vicinity of a mineral grain may cause a sequence of reactions to occur such that phlogopite first alters to chlorite, which changes to Mg-montmorillonite, which ultimately decomposes to kaolinite or even gibbsite. Similarly, illite may form either microcline or muscovite by a reduction of $[K^+]$ or decrease in pH. Depending on the existence of the appropriate chemical gradients, muscovite flakes may alter to kaolinite or illite, which may transform into illite-montmorillonite or illite-kaolinite mixed-layer clays. However, predictions of reaction products based on thermodynamics alone are not always reliable because unfavorable kinetics may prevent the reactions from reaching equilibrium or from going to completion.

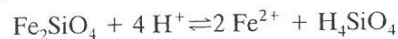
b. Solubility of Ferrous Silicates

The ferrous silicates include fayalite (Fe_2SiO_4), ferrosilite ($FeSiO_3$), and the two-layer phyllosilicate *greenalite*. In addition, the carbonate mineral siderite ($FeCO_3$) and the hydroxide $[Fe(OH)_2]$ may form. We assume that Fe is predominantly in the divalent state and that minerals containing Fe^{3+} do not occur in the system.

The ferrous silicates dissolve congruently much like the Mg-silicates we discussed in Chapter 12 (Figure 12.5). The solution of the ferrous silicates is expressed by the equation in Table 13.7, which have been plotted in Figure 13.4 in coordinates of $\log [Fe^{2+}]/[H^+]^2$ and $\log [H_4SiO_4]$. The lines are the boundaries between solutions that are undersaturated or supersaturated with respect to each mineral. For example, fayalite dissolves in solutions represented by points lying to the left of the equilibrium line (equation 1, Table 13.7). Evidently, greenalite (equation 3, Table 13.7) is less soluble, and therefore more stable, than either fayalite or ferrosilite in contact with water at 25°C. However, the composition of natural solutions in contact with the atmosphere is restricted by the solubility boundaries of siderite (equation 4, Table 13.7) and amorphous silica (equation 6, Table 13.7). Therefore, even greenalite dissolves in natural solutions and precipitates only from solutions that have low bicarbonate ion activities.

Table 13.7 Solubility Equations of Pure Fe-Silicate Minerals in Pure Water at 25°C^a

Fayalite



$$\Delta G_R^\circ = -20.76 \text{ kcal} \quad K = 10^{15.22}$$

$$\log \frac{[Fe^{2+}]}{[H^+]^2} = -0.5 \log [H_4SiO_4] + 7.61 \quad (1)$$

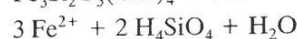
Ferrosilite



$$\Delta G_R^\circ = -7.663 \text{ kcal} \quad K = 10^{5.62}$$

$$\log \frac{[Fe^{2+}]}{[H^+]^2} = -\log [H_4SiO_4] + 5.62 \quad (2)$$

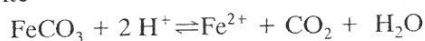
Greenalite



$$\Delta G_R^\circ = -18.577 \text{ kcal} \quad K = 10^{13.60}$$

$$\log \frac{[Fe^{2+}]}{[H^+]^2} = -0.67 \log [H_4SiO_4] + 0.53 \quad (3)$$

Siderite



$$\Delta G_R^\circ = +0.509 \text{ kcal} \quad K = 10^{-0.37}$$

$$\log \frac{[Fe^{2+}]}{[H^+]^2} = 3.15 \text{ at } [CO_2] = 3 \times 10^{-4} \text{ atm} \quad (4)$$

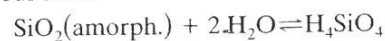
Ferrous hydroxide



$$\Delta G_R^\circ = -15.924 \text{ kcal} \quad K = 10^{11.67}$$

$$\log \frac{[Fe^{2+}]}{[H^+]^2} = 11.67 \quad (5)$$

Amorphous silica



$$\Delta G_R^\circ = +4.044 \text{ kcal} \quad K = 10^{-2.96}$$

$$\log [H_4SiO_4] = -2.96 \quad (6)$$

^aBased on thermodynamic constants in Appendix B.

The solubility data in Figure 13.4 indicate that greenalite could precipitate in place of siderite if $\log [Fe^{2+}]/[H^+]^2$ is greater than about 7. The fugacity of CO_2 that would permit this can be calculated from the solubility equation for siderite in Table 13.7:

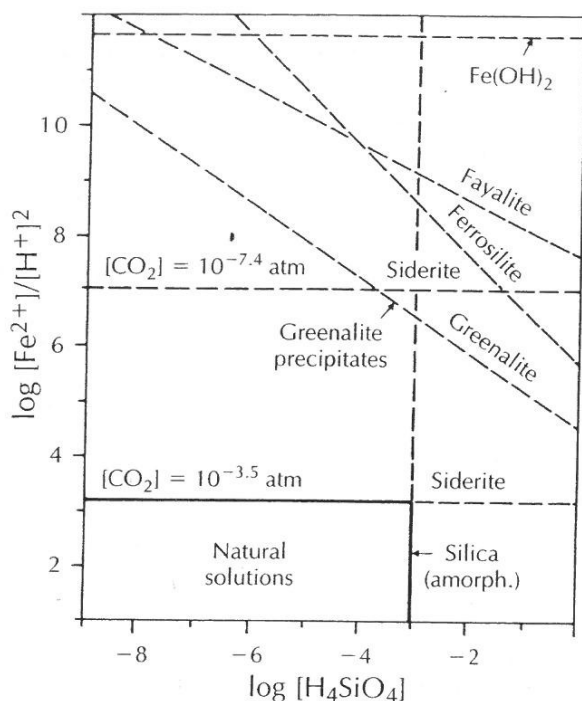


Figure 13.4 Solubility of ferrous silicates in pure water at 25°C. Greenalite, which is a two-layer clay analogous to kaolinite and serpentine, is soluble in natural solutions open to the atmosphere. It can precipitate at 25°C, only when $[\text{CO}_2] = 10^{-7.4}$ atm or less, from solutions that are nearly saturated with respect to amorphous silica. Greenalite is the most abundant primary Fe-silicate in the Gunflint Iron Formation of Proterozoic age in western Ontario, Canada, and its absence from the siderite facies implies that the siderite facies was deposited where the fugacity of CO_2 was greater than $10^{-7.4}$ atm.

$$\frac{[\text{Fe}^{2+}]}{[\text{H}^+]^2} = \frac{10^{-0.373}}{[\text{CO}_2]} = 10^{7.0} \quad (13.35)$$

which yields $[\text{CO}_2] = 10^{-7.4}$ atm. In other words, greenalite can precipitate only from solutions that have very low carbonate ion concentrations and are close to being saturated with respect to amorphous silica. Such solutions may be discharged by volcanic hotsprings in shallow marine basins where cherty iron formations were deposited in Precambrian time. For example, Goodwin (1956) reported that greenalite is the

most common iron silicate mineral in unaltered parts of the Gunflint Formation in western Ontario, Canada. It occurs as light green to olive green granules up to 3 mm in diameter, except in the banded chert-siderite and tuffaceous shale facies. The absence of greenalite from siderite-bearing rocks is explained by the difference in their solubilities when the solutions are in contact with CO_2 of the atmosphere. Hydrothermal solutions discharged into the Gunflint basin could precipitate greenalite and amorphous silica if they had sufficiently low carbonate ion concentrations to prevent siderite from forming.

Minnesotaite is common in the Fe-silicate facies of the Biwabik Iron Formation of northern Minnesota, but is absent from the Gunflint Formation. James (1966) pointed out that this difference between the two iron formations may be related to the fact that the Biwabik Formation was intruded by the Embarras Granite over much of its length, whereas the Gunflint Formation was not intruded by granite. Consequently, he agreed with Tyler (1949) that minnesotaite is a product of metamorphism, which made the Biwabik Formation more susceptible to chemical weathering than the Gunflint Formation and resulted in the formation of hematite ore bodies in Minnesota.

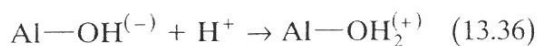
13.5 Colloidal Suspensions and Ion Exchange

When clay minerals and other small particles with diameters of $<1 \mu\text{m}$ are dispersed in pure water, they form a colloidal suspension called a *sol*. Suspensions of small droplets of a liquid that is immiscible in another liquid are known as *emulsions*, and solid particles or liquid droplets dispersed in a gaseous medium form *aerosols*. Such colloidal suspensions may appear to be homogeneous, and the particles do not settle out under the influence of gravity. Some kinds of particles in sols attract water molecules to themselves and are therefore said to be *hydrophilic*, whereas others lack this tendency and are classified as *hydrophobic*. A sol composed of hydrophilic particles may *set* to form a translucent or transparent *gel*. Some

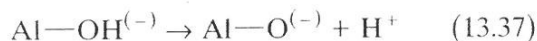
gels can be dispersed again into a sol, whereas in others the process of gel formation is irreversible. For example, concentrated solutions of silic acid can polymerize to form amorphous silica particles, and the resulting colloidal suspension can set to form a gel. However, the silica gel does not redisperse spontaneously, but slowly loses water as it crystallizes to form quartz (variety chalcedony) with opal as an intermediate product (see Section 9.6).

The properties of colloidal suspensions arise from electrical charges on the surface of the particles. The presence of electrical charges on colloidal particles is related to their small size and resultant large surface area. The atoms on the surfaces of solid particles of all kinds of materials are incompletely bonded because they are not surrounded by ions of the opposite charge. Therefore, such particles attract ions from the solution in which they are suspended and thereby acquire a characteristic charge. When *clay minerals* are suspended in water, some of the adsorbed interlayer ions may be lost, leaving the individual particles with a negatively charged surface. In addition, clay particles also have surface ions with unsatisfied charges and hence become charged not only because of internal crystallographic reasons but also because of the surface effects that occur in all small particles.

The surfaces of silicate, oxide, or hydroxide particles expose atoms of Si, Al, Fe, Mn, etc., which are bonded to O^{2-} or OH^- radicals within the solid. However, OH^- radicals on the surface exert an attractive force on positively charged ions in the solution, including H^+ . Therefore, such particles acquire a *positive* charge in strongly acidic solutions by virtue of the reaction:



where the parentheses indicate "net" positive or negative charges that are not necessarily equal to one electronic unit. In strongly basic solutions the OH^- radicals may release H^+ , thereby causing a net *negative* charge at the surface:



Similar reactions may take place with atoms of Si, Fe, Mn, etc., on the surface of colloidal particles that contribute to their total electrical charges.

The presence of electrical charges on the surface of colloidal particles overcomes their tendency to settle out and keeps them in suspension indefinitely unless the charges are neutralized by the addition of an electrolyte or the passage of an electrical current. In general, colloidal suspensions are stabilized in dilute electrolyte solutions because ions from the solution are needed to build up the surface charge. However, at high electrolyte concentrations the surface charges of the particles are neutralized and the suspension *coagulates*, that is, the particles settle out.

The polarity of the surface charge of colloidal particles depends on the activity of H^+ in the water. Consequently, we define the *isoelectric point* as the pH of the aqueous medium at which the surface charge of particles of a specific material is equal to zero. If the pH is less than the isoelectric point, the charge is *positive* by virtue of equation 13.36; whereas if the pH is greater than the isoelectric points the charge is *negative* according to equation 13.37. The isoelectric points of different kinds of materials in Table 13.8 range from pH = 1.0 (amorphous silica) to pH = 9.1 (Al_2O_3 , corundum). The isoelectric points can be used to deduce the polarity of the surface charge of colloidal particles of each material in acid and basic media. The data in Table 13.8 indicate that colloidal particles composed of amorphous silica, quartz, Mn-dioxide, montmorillonite, and kaolinite have a *negative* surface charge, except in highly acidic environments caused, for example, by local discharge of acid mine waters. The oxides of Fe form particles whose surface charges may change from positive to negative between pH values 5 and 7. However, Fe-oxide or hydroxide and TiO_2 particles in the oceans (pH = 8.2) have *negative* charges. The oxides and hydroxides of Al stand out from the others in Table 13.8 because they form *positively* charged particles in all but highly basic environments.

It turns out that the presence of electrolytes in a colloidal suspension affects the development

Table 13.8 Isoelectric Points of Various Natural Materials When Suspended in Water of Varying Acidity but Lacking Dissolved Electrolytes

<i>Material</i>	<i>pH</i>	<i>Polarity in acid medium</i>	<i>Polarity in basic medium</i>
SiO ₂ (amorph.)	1.0–2.5	– ^a	–
Quartz	2.0	–	–
MnO ₂ (different forms)	2.0–4.5	–	–
Albite	2.0	–	–
Montmorillonite	2.5	–	–
Kaolinite	4.6	–	–
Hematite	5–9	±	±
	commonly 6–7		
Magnetite	6.5	+	–
Goethite	6–7	±	±
Limonite (Fe ₂ O ₃ · <i>n</i> H ₂ O)	6–9	±	±
Anatase (TiO ₂)	7.2	+	–
Gibbsite	~9	+	+
Corundum	9.1	+	+
Periclase (MgO)	12.4	+	+

^aThe polarity changes from negative to positive at pH values less than the isoelectric point. Thus, the charge is negative if the pH is greater than the isoelectric point. *Acid* and *basic* are defined relative to pH = 7.0 at 25°C.

After Drever (1982), Table 4.2; Stumm and Morgan (1970).

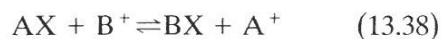
of surface charges of the particles. Therefore, the pH at which surface charges of colloidal particles become equal to zero in the presence of electrolytes differs from the isoelectric point in ways that depend on the composition of the particles and the charges of the ions of the electrolyte. The pH at which the surface charge of colloidal particles in the presence of electrolytes is equal to zero is known as the *zero point of charge (ZPC)*. The numerical values of the ZPC depend not only on the composition of the colloidal particles but also on the composition of the electrolytes in the suspension.

The charges of colloidal particles of different compositions affect geochemical processes in the oceans, in soils, and in the deposition of sediment because negatively charged particles, of colloidal

size or larger, adsorb cations (metals), whereas positively charged particles adsorb anions (non-metals). Particles of all kinds that sink in the oceans scavenge metal ions from seawater and transfer them to the sediment accumulating at the bottom of the oceans. Consequently, deep-sea sediment is enriched in many metals compared to normal marine sediment, which is dominated by detrital sediment derived from nearby continental areas (Section 4.5).

Colloidal particles in soils act as temporary repositories of ions released by chemical weathering or by the addition of fertilizers. The resulting capacity of soils to store plant nutrients enhances their fertility. In addition, the fixation of ions in soils and sediment enhances the solubility of minerals and results in the purification of groundwater.

When clay minerals or other particles of colloidal or larger size are dispersed in water, their surfaces become charged and attract ions of the opposite charge, which are adsorbed by means of electrostatic bonds. With a few exceptions, mineral surfaces are not particularly selective and attract ions of different elements without much discrimination. This means that the adsorbed ions are *exchangeable* for others in the solution depending primarily on their activity. The resulting phenomenon of *ion exchange* can be represented by a simple exchange reaction:



which indicated that ion A^+ on the adsorber is exchanged for ion B^+ in the solution. Such exchange reactions reach equilibrium very rapidly and are then subject to the Law of Mass Action; that is:

$$\frac{[BX][A^+]}{[AX][B^+]} = K_{ex} \quad (13.39)$$

where K_{ex} is the exchange constant. If the exchanger accepts A^+ and B^+ with equal ease, then $K_{ex} = 1.0$. However, in general this is not the case because A^+ and B^+ do not have the same charge/radius ratio and their hydrated ions have different radii. Moreover, divalent or trivalent ions form stronger electrostatic bonds than univalent ions and may therefore be adsorbed preferentially.

Equation 13.39 can be restated in the form:

$$\frac{[A^+]}{[B^+]} = K_{ex} \left(\frac{(AX)}{(BX)} \right)^n \quad (13.40)$$

where (AX) and (BX) are the molar concentrations of A and B on the exchanger and n is an exponent with values between 0.8 and 2.0 (Berner, 1971; Garrels and Christ, 1965). These equations explain why the activities of ions in the aqueous phase of colloidal suspension are controlled by ion exchange. For example, if A^+ is added to a colloidal solution containing A^+ and B^+ in ion exchange equilibrium with clay minerals in suspension, A^+ will displace B^+ from exchange sites and thereby reestablish equilibrium in accordance with Le Châtelier's principle.

The capacity of clay minerals to act as ion exchangers varies widely, depending at least in part on the charge imbalances that originate within the crystal lattice. For example, kaolinite commonly has a low *cation exchange capacity* (*CEC*), whereas montmorillonite and vermiculite have large CECs. The CEC of a sample of clay is determined by means of a standard procedure: 100 g of dry clay is repeatedly placed into NaCl solutions at pH = 7.0 until all exchangeable cations are replaced by Na^+ . The Na-saturated clay is then added to concentrated KCl solutions and the amount of Na^+ released by the clay at pH = 7.0 is determined by analysis and is expressed in units of milliequivalents per 100 g of dry clay. The CEC values of clay minerals are related to their chemical compositions and therefore vary widely. However, the values in Table 13.9 are fairly typical of the respective clay minerals.

The CEC of a clay mineral can be estimated from the net negative charge of its structural formula. For example, the montmorillonite represented by formula 13.3 has a charge of -0.46 electronic units per formula weight which is neutralized by exchangeable Ca^{2+} , Mg^{2+} , and Na^+ . The CEC of this mineral caused only by the charge in the layers is:

Table 13.9 Cation Exchange Capacity (CEC) of Various Clay Minerals

Mineral	CEC, meq/100 g at pH = 7.0
Kaolinite	3–15
Chlorite	10–40?
Illite	10–40
Glauconite	11–20+
Palygorskite	20–30
Allophane	~70
Smectite (montmorillonite)	70–100
Vermiculite	100–150

SOURCE: Garrels and Christ (1965).

$$\text{CEC} = \frac{0.46 \times 100 \times 1000}{360.25} = 127.7 \text{ meq/100 g} \quad (13.41)$$

This value is probably an underestimate because it does not include charges caused by broken bonds. Nevertheless, the result indicates that 100 g of this clay can adsorb 64.6 mmol of Ca^{2+} , which amounts to $40.08 \times 64.6/1000 = 2.59$ g of Ca or 2.97 g of Na. These amounts are not only substantial, but the presence of Na^+ , Ca^{2+} , or other cations has a significant effect on the mechanical properties of the clay.

The presence of Na^+ in exchangeable positions increases the plasticity of clay and can make soil containing such clay unsuitable for agriculture. The best soils contain clay minerals having a mixture of Ca^{2+} , H^+ , Mg^{2+} , and K^+ as exchangeable cations. However, in the ceramic industry, where plasticity is a desirable property, clay may be converted to the Na-form by suitable treatment prior to its use.

Smectite clays expand when water molecules replace exchangeable ions, and slurries of smectite may set to form a solid that can revert to a liquid when it is disturbed. This property, called *thixotropy*, characterizes bentonite, which is composed largely of smectite clay minerals, and makes it useful as a sealant of building foundations and as a drilling mud in the petroleum industry.

13.6 Dating of Clay Minerals

One of the primary tasks in the Earth Sciences is to establish an accurate chronology of the history of the Earth, which is recorded in the sedimentary rocks that have survived to the Present. This has turned out to be a very difficult task because sedimentary rocks are generally not datable by methods based on radioactivity. One of several reasons for this difficulty is that sedimentary rocks are composed primarily of detrital particles of minerals that crystallized to form igneous or metamorphic rocks long before they were incorporated into clastic sedimentary rocks, such as sandstones or shales. One can make a virtue of

this circumstance by dating detrital minerals such as muscovite, feldspar, or zircon to determine their *provenance* by identifying their source by its age (Faure and Taylor, 1981). However, the age of sedimentary rocks can be determined only by dating minerals that formed at the time of deposition (Clauer and Chaudhuri, 1995).

Unfortunately, most of the *authigenic* minerals in sedimentary rocks are not suitable for dating by isotopic methods because they do not contain radioactive parent elements in sufficient concentration or because they do not retain the radiogenic daughter isotope quantitatively. These shortcomings effectively eliminate several authigenic minerals that commonly occur in sedimentary rocks, including calcite, dolomite, quartz, gypsum, anhydrite, hematite, magnetite, pyrite, pyrolusite, halite, sylvite, and other evaporite minerals. However, certain clay minerals are authigenic and do contain K and Rb, which make them potentially datable by the K–Ar and Rb–Sr methods described in Chapter 16.

Glaucinite and illite are best suited for dating of sedimentary rocks, but only under narrowly defined circumstances. For example, glaucony in the form of greenish pellets and encrustations in marine sedimentary rocks contains varying amounts of K, which stabilizes the lattice by strengthening the bonds between clay units. As a result, the retentivity of glaucony with respect to radiogenic ^{40}Ar improves with increasing K content and the mineral becomes closed to ^{40}Ar when $\text{K}_2\text{O} > 6.0\%$ (Odin and Matter, 1981). Therefore, only glauconitic micas with the required K concentration can yield reliable dates by the K–Ar method. The selection criteria for dating glaucony by the Rb–Sr method are not as well defined, partly because radiogenic ^{87}Sr is less mobile than ^{40}Ar . However, both methods are also sensitive to gain or loss of the radioactive parent element so that well-crystallized glauconitic mica is also preferable for dating by the Rb–Sr method. However, even well-crystallized glauconies may be appreciably altered by thermal metamorphism, tectonic deformation, and

ion exchange with brines in the subsurface. Therefore, glauconies that remember their age are valuable exceptions rather than the rule (Odin, 1982).

The application of the Rb–Sr method to dating *clay-rich* sedimentary rocks has been summarized by Clauer (1982). The whole-rock isochron method sometimes yields useful results (Cordani et al., 1985), especially in unfossiliferous shales of Precambrian age containing large amounts of accumulated radiogenic ^{87}Sr . In general, however, it is better to use methods that rely on the properties of a specific authigenic mineral, such as illite. Morton (1985) demonstrated that coarse-grained illite (1–2 μm) in Upper Devonian black shales of Texas is composed of the 2M polymorph of detrital origin and yielded a “provenance date” by the Rb–Sr method that predates the time of deposition. However, 1Md illite in the $<0.2\text{-}\mu\text{m}$ fraction in this case dated the time of diagenetic alteration of the illite by migrating brines. The fit of data points to a Rb–Sr isochron was dramatically improved by treating fine-grained 1Md illites with ammonium acetate, which removed Rb^+ and Sr^{2+} from exchangeable sites.

13.7 Summary

Clay minerals are classified according to their crystal structure into platy clays, fibrous clays, and x-ray amorphous clays. The platy clays are composed of sheets of silica tetrahedra joined at their bases (tetrahedral layer, T) and Al hydroxide octahedra joined by sharing hydroxyl ions at the corners (octahedral layer, O). The phyllosilicate clays are made up of two-layer (T–O) and three-layer (T–O–T) packages. The chemical diversity of clay minerals results from complete or partial replacement of Al^{3+} in the octahedral layer and of Si^{4+} in the tetrahedral layer.

In the octahedral layer, Al^{3+} can be replaced by Fe^{3+} , Fe^{2+} , Mg^{2+} , Zn^{2+} , Li^+ , and Cr^{3+} , whereas

in the tetrahedral layer Si^{4+} is partially replaced by Al^{3+} . The charge imbalance that results from this substitution in the tetrahedral layer may be partly compensated by additional cations in the octahedral layer. However, the net negative charge arising from either the tetrahedral or the octahedral layer, or both layers, is neutralized by cations that are adsorbed to the outer surfaces of the tetrahedral layer and thereby bond individual clay units to each other. Potassium ions are especially effective in this regard because they fit into the hexagonal rings of silica tetrahedra that make up the tetrahedral layer. Consequently, K-bearing clays like illite have short interlayer spacings and do not expand in water. Clays that have adsorbed Na^+ , Ca^{2+} , or Mg^{2+} ions have wider interlayer spacings than K-clays and expand in water because the interlayer cations are displaced by water molecules.

The chemical diversity of clay minerals makes it difficult to ascertain their thermodynamic properties. Several empirical schemes have been proposed to overcome this difficulty, including one by Tardy and Garrels (1974). Their method for estimating standard free energies is based on the oxides and hydroxides that make up the layers of two- or three-layer clays and on the adsorbed ions, including H^+ . However, in order to construct mineral stability diagrams of general validity, standardized smectite-type clay minerals are used. The results indicate that primary mica minerals (muscovite, biotite, and phlogopite) alter to a variety of clay minerals depending on environmental conditions. Mineral stability diagrams can be used to understand natural geochemical processes such as the formation of greenalite in cherty iron formations and the incompatibility of greenalite and siderite.

Clay minerals and other small particles can form colloidal suspensions in dilute electrolyte solutions because of electrical charges on their surfaces. Such suspensions may be stable for long periods of time but are coagulated by an increase in the concentration of electrolytes. The polarity of the electrical charges of colloidal particles of different compositions depends on the pH of the

solution. Clay minerals, quartz, feldspar, and Mn oxides are negatively charged except in strongly acidic environments with $\text{pH} < 2.0$. The various oxides and hydroxides of Fe can have positive or negative surface charges in the pH range from 5 to 9, whereas Al oxides and hydroxide particles have positive charges except at $\text{pH} > 9$.

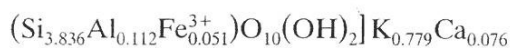
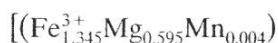
Cations that are adsorbed on the charged surfaces of small particles in contact with dilute electrolyte solutions are exchangeable for ions in the solution. These cation exchange reactions can achieve a state of equilibrium that causes water-particle mixtures to respond to changes in water compositions in accordance with Le Châtelier's principle. The ability of small particles of different minerals to adsorb exchangeable ions is expressed by its cation exchange

capacity (CEC), which can be measured by means of a standardized procedure. The physical properties of clays are strongly affected by the kinds of adsorbed cations.

Certain authigenic K-bearing clay minerals are dateable by the K-Ar and Rb-Sr methods and therefore provide rare opportunities to determine the ages of sedimentary rocks. However, only clay samples whose chemical composition, crystallinity, or grain size has been evaluated should be used, and the resulting dates may be affected by thermal metamorphism, structural deformation, or ion exchange with circulating brines. Dating of sedimentary rocks by means of clay minerals requires a thorough understanding of clay mineralogy as well as of geochemistry.

Problems

1. Determine the electrical charges of the octahedral and tetrahedral layers and check whether the exchangeable ions balance the charge in the following clay mineral.



2. A glauconite from Point Jackson, Francosia, Wisconsin, has the following chemical composition (Johnston and Cardile, 1987).

Element	Moles	Element	Moles
Si	3.611	Mn	0.001
Al	1.238	Ca	0.096
Fe^{2+}	1.097	K	0.725
Mg	0.442	H	2.000
Ti	0.003	O	12.000

Rewrite this analysis in the form of a structural formula for a three-layer clay and test it for electrical neutrality.

3. Recalculate the chemical analysis of an impure clay mineral from Macon, Georgia, into a structural formula for a two-layer clay (Grim, 1968).

Component	%	Component	%
SiO_2	45.20	K_2O	0.49
Al_2O_3	37.02	Na_2O	0.36
Fe_2O_3	0.27	TiO_2	1.26
FeO	0.06	$\text{H}_2\text{O}-$	1.55 ^a
MgO	0.47	$\text{H}_2\text{O}+$	13.27
CaO	0.52	Sum	100.47

^a $\text{H}_2\text{O}-$ is water removed at 105°C , whereas $\text{H}_2\text{O}+$ is lost only at high temperature.

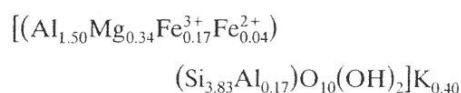
4. Recalculate the chemical analysis of a clay mineral from Montmorillon, France, into a structural formula for a three-layer clay (Grim, 1968).

Component	%	Component	%
SiO_2	51.14	K_2O	0.11
Al_2O_3	19.76	Na_2O	0.04
Fe_2O_3	0.83	TiO_2	—
FeO	—	$\text{H}_2\text{O}-$	14.81
MgO	3.22	$\text{H}_2\text{O}+$	7.99
CaO	1.62	Sum	99.52

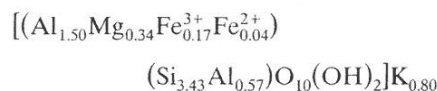
5. Recalculate the chemical analysis of a chlorite mineral from Schmiedefeld, Thuringia, into a structural formula (Grim, 1968).

Component	%	Component	%
SiO ₂	20.82	MgO	4.15
Al ₂ O ₃	17.64	H ₂ O+	10.30
Fe ₂ O ₃	8.70	Sum	99.57
FeO	37.96		

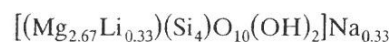
6. Estimate the standard free energy of formation of "average" montmorillonite by the method of Tardy and Garrels (1974) based on the formula



7. Estimate the standard free energy of formation of "pure" illite by the method of Tardy and Garrels (1974) based on the formula



8. Calculate the solubility product constant for hectorite at 25°C, assuming congruent solution based on the formula



References

- BERNER, R. A., 1971. *Principles of Chemical Sedimentology*. McGraw-Hill, New York, 240 pp.
- BRADLEY, W. F., 1940. The structural scheme of attapulgite. *Amer. Mineral.*, **25**:405-410.
- BRINDLEY, G. S., 1961. Kaolin, serpentine, and kindred minerals. In G. Brown (Ed.), *The X-ray Identification and Crystal Structures of Clay Minerals*, 2nd ed., 51-131. Mineralogy Society London, Monograph.
- BRINDLEY, G. S., S. W. BAILEY, G. T. FAUST, S. A. FORMAN, and C. I. RICH, 1968. Report of the Nomenclature Committee (1966-67) of the Clay Minerals Society. *Clays and Clay Minerals*, **16**:322-324.
- BROWN, B. E., and S. W. BAILEY, 1962. Chlorite polytypisms, part I, regular and semi-random one-layer structures. *Amer. Mineral.*, **47**:819-850.
- BURST, J. F., 1958. Glauconite pellets, their mineral nature and application to stratigraphic interpretations. *Amer. Assoc. Petrol. Geologists Bull.*, **42**:310-327.
- CAIRNS-SMITH, A. G., and H. HARTMAN (Eds.), 1987. *Clay Minerals and the Origin of Life*. Cambridge University Press, Cambridge, England, 202 pp.
- CARROLL, D., 1970. Clay minerals: a guide to their x-ray identification. *Geol. Soc. Amer., Special Paper* 126, 80 pp.
- CHEN, C.-H., 1975. A method of estimating standard free energies of formation of silicate minerals at 298.15 K. *Amer. J. Sci.*, **275**:801-817.
- CLAUER, N., 1982. The rubidium-strontium method applied to sediments: certitudes and uncertainties. In G. S. Odin (Ed.), *Numerical Dating in Stratigraphy*, vol 1, 245-276. Wiley, Chichester, England, 630 pp.
- CLAUER, N., and S. CHAUDHURI, 1995. *Clays in Crustal Environments; Isotopic Dating and Tracing*. Springer-Verlag, Heidelberg, 359 pp.
- CORBATO, C. E., and R. T. TETTENHORST, 1987. Analysis of illite-smectite interstratification. *Clay Miner.*, **22**:269-285.
- CORDANI, U. G., A. THOMAZ-FILHO, B. B. BRITO-NEVES, and K. KAWASHITA, 1985. On the applicability of the Rb-Sr method to argillaceous sedimentary rocks: some examples from Precambrian sequences of Brazil. *Giornale di Geologia, ser. 3a.*, **47**:253-280.
- DREVER, J. I., 1982. *The Geochemistry of Natural Waters*. Prentice-Hall, Upper Saddle River, NJ, 388 pp.
- EGGLESTON, R. A., and S. W. BAILEY, 1967. Structural aspects of dioctahedral chlorite. *Amer. Mineral.*, **52**:673-689.
- FAURE, G., and K. S. TAYLOR, 1981. Provenance of some glacial deposits in the Transantarctic Mountains based on Rb-Sr dating of feldspars. *Chem. Geol.*, **32**:271-290.
- FOSTER, M. D., 1962. Interpretation of the composition of trioctahedral micas. *U. S. Geol. Surv. Prof. Paper* **354-B**:11-49.
- GARRELS, R. M., and C. L. CHRIST, 1965. *Solutions, Minerals, and Equilibria*, Harper & Row, New York, 450 pp.
- GOODWIN, A. M., 1956. Facies relations in the Gunflint Iron Formation. *Econ. Geol.*, **51**:588-595.
- GRIM, R. E., 1968. *Clay Mineralogy*, 2nd ed. McGraw-Hill, New York, 596 pp.

- GRIM, R. E., R. H. BRAY, and W. F. BRADLEY, 1937. The mica in argillaceous sediments. *Amer. Mineral.*, **22**:813–827.
- GRUNER, J. W., 1934. The structure of vermiculites and their collapse by dehydration. *Amer. Mineral.*, **19**:557–575.
- HELGESON, H. C., 1969. Thermodynamics of hydrothermal systems at elevated temperatures and pressures. *Amer. J. Sci.*, **267**:729–804.
- HELGESON, H. C., J. M. DELANY, H. W. NESBITT, and D. K. BIRD, 1978. Summary and critique of the thermodynamic properties of rock-forming minerals. *Amer. J. Sci.*, **278A**:1–229.
- HEY, M. H., 1954. New review of the chlorites. *Mineral. Mag.*, **30**:277–292.
- HOWER, J., 1961. Some factors concerning the nature and origin of glauconite. *Amer. Mineral.*, **46**:313–334.
- JAMES, H. L., 1966. Chemistry of the iron-rich sedimentary rocks. In M. L. Fleischer (Ed.), *Data of geochemistry. U. S. Geol. Surv. Prof. Paper*, 440-W; 60 pp.
- JOHNSTON, J. H., and C. M. CARDILE, 1987. Iron substitution in montmorillonite, illite, and glauconite by ^{57}Fe Mössbauer spectroscopy. *Clays and Clay Minerals*, **35**:170–176.
- KITTRICK, J. A., 1971a. Stability of montmorillonites: I. Belle Fourche and Clay Spur montmorillonites. *Soil Sci. Soc. Amer. Proc.*, **35**:140–145.
- KITTRICK, J. A., 1971b. Stability of montmorillonites: II. Aberdeen montmorillonite. *Soil Sci. Soc. Amer. Proc.*, **35**:820–823.
- LINDSAY, W. L., 1979. *Chemical Equilibria in Soils*. Wiley, New York, 449 pp.
- MCRAE, S. G., 1972. Glauconite. *Earth Sci. Rev.*, **8**(4):397–440.
- MILLOT, G., 1970. *Geology of Clays*. Springer-Verlag, Berlin, 429 pp.
- MORTON, J. P., 1985. Rb–Sr dating of diagenesis and source age of clays in Upper Devonian black shales of Texas. *Geol. Soc. Amer. Bull.*, **96**:1043–1049.
- NRIAGU, J. O., 1975. Thermochemical approximations for clay minerals. *Amer. Mineral.*, **60**:834–39.
- ODIN, G. S. (Ed.), 1982. *Numerical Dating in Stratigraphy*, vols. 1 and 2. Wiley, Chichester, England, 1040 pp.
- ODIN, G. S., and A. MATTER, 1981. De glauconiarum origine. *Sedimentology*, **28**:611–641.
- REYNOLDS, R. C., 1967. Interstratified clay systems—calculation of the total one-dimensional diffraction function. *Amer. Mineral.*, **52**:661–672.
- REYNOLDS, R. C., 1980. Interstratified clay minerals. In G. W. Brindley and G. Brown (Eds.), *Crystal Structures of Clay Minerals and Their X-ray Identification*. Mineral. Soc. London, Monograph 5, 495 pp.
- ROUTSON, R. C., and J. A. KITTRICK, 1971. Illite solubility. *Soil Sci. Soc. Amer. Proc.*, **35**:714–718.
- SINGER, A., and E. GALAN (Eds.), 1984. *Palygorskite-Sepiolite. Occurrences, Genesis and Uses*. Elsevier, Amsterdam, 352 pp.
- SMITH, J. V., and H. S. YODER, 1956. Studies of mica polymorphs. *Mineral. Mag.*, **31**:209–235.
- STUMM, W., and J. J. MORGAN, 1970. *Aquatic Chemistry*. Wiley, New York, 583 pp.
- SUDO, T., and S. SHIMODA (Eds.), 1978. *Clays and Clay Minerals in Japan*. Elsevier, Amsterdam.
- SUDO, T., S. SHIMODA, H. YOTSUMOTO, and S. AITA, 1981. *Electron Micrographs of Clay Minerals*. Elsevier, Amsterdam, 204 pp.
- TARDY, Y., and R. M. GARRELS, 1974. A method of estimating the Gibbs energies of formation of layer silicates. *Geochim. Cosmochim. Acta*, **38**:1101–1116.
- TARDY, Y., and R. M. GARRELS, 1976. Prediction of Gibbs energies of formation—I. Relationships among Gibbs energies of formation of hydroxides, oxides and aqueous ions. *Geochim. Cosmochim. Acta*, **40**:1051–1056.
- TARDY, Y., and R. M. GARRELS, 1977. Prediction of Gibbs energies of formation of compounds from the elements—II. Monovalent and divalent metal silicates. *Geochim. Cosmochim. Acta*, **41**:87–92.
- TARDY, Y., and L. GARTNER, 1977. Relationships among Gibbs energies of formation of sulfates, nitrates, carbonates, oxide, and aqueous ions. *Contrib. Mineral. Petrol.*, **63**:89–102.
- TARDY, Y., and P. VIEILLARD, 1977. Relationships among Gibbs free energies and enthalpies of formation of phosphates, oxides and aqueous ions. *Contrib. Mineral. Petrol.*, **63**:75–88.
- TARDY, Y., 1979. Relationships among Gibbs energies of formation of compounds. *Amer. J. Sci.*, **279**:217–224.
- TYLER, S. A., 1949. Development of Lake Superior soft iron ores from metamorphosed iron formation. *Geol. Soc. Amer. Bull.*, **60**:1101–1124.
- VELDE, B., 1977. *Clays and Clay Minerals in Natural and Synthetic Systems*. Elsevier, Amsterdam, 218 pp.
- WEAVER, C. E., 1956. The distribution and identification of mixed-layer clays in sedimentary rocks. *Amer. Mineral.*, **41**:202–221.
- WEAVER, R. M., M. L. JACKSON, and J. K. SYERS, 1971. Magnesium and silicon activities in matrix solutions of montmorillonite-containing soils in relation to clay mineral stability. *Soil Sci. Soc. Amer. Proc.*, **35**:823–830.
- WEAVER, C. E., and L. D. POLLARD, 1973. *The Chemistry of Clay Minerals*. Elsevier, Amsterdam.
- YODER, H. S., and H. P. EUGSTER, 1955. Synthetic and natural muscovites. *Geochim. Cosmochim. Acta*, **8**:225–280.
- ZEN, E.-AN, 1967. Mixed-layer minerals as one-dimensional crystals. *Amer. Mineral.*, **52**:635–660.
- ZEN, E.-AN, 1972. Gibbs free energy, enthalpy and entropy of ten rock-forming minerals: calculations, discrepancies, implications. *Amer. Mineral.*, **57**:524–553.

Oxidation–Reduction Reactions

We now return to a subject we first considered in Chapter 7 and mentioned again in connection with fugacity diagrams in Section 12.5. In Chapter 7 we divided the elements of the periodic table into the *electron donors* (metals) and *electron acceptors* (nonmetals) based on their electron configurations. Metals are electron donors because their available orbitals are largely empty, whereas those of nonmetals are almost filled to capacity. As a result, the atoms of metals give up electrons from partially filled orbitals to form *cations*, whereas atoms of nonmetals form *anions* by attracting electrons to themselves until their available orbitals are filled. These tendencies of the elements are expressed numerically by means of the *electronegativity*, which is defined as the ability of an atom in a molecule to attract electrons to itself.

Some elements always give up or accept the same number of electrons and therefore have fixed valences. For example, the alkali metals and alkaline earths give up one or two electrons, respectively, whereas the halogens accept one electron under natural conditions. Other elements, such as Fe, Cu, and U can lose varying numbers of electrons depending on the strength of the electron acceptor they are interacting with. Similarly, nonmetals such as S, N, and C can take up varying numbers of electrons depending on their availability in the environment.

These concepts come into play in chemical reactions in which the valences of some of the participating elements change. The loss of electrons by the electron donor is called *oxidation* and the gain of electrons by the electron acceptor is called *reduction*. Oxidation must be accompanied by

reduction because the number of electrons released by the donor must equal the number of electrons gained by the acceptor. As a result of the transfer of electrons in an oxidation–reduction reaction the electron donor is *oxidized* and the electron acceptor is *reduced*. Evidently, the electron donor is the *reducing agent* and the electron acceptor is the *oxidizing agent*. Since the atoms of metals are electron donors, we can say that metals are reducing agents and that nonmetals are oxidizing agents.

14.1 Balancing Equations of Oxidation–Reduction Reactions

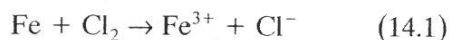
In order to balance oxidation–reduction reactions correctly, it is helpful to make use of the *valence number* (also known as the oxidation number), which is defined as the electrical charge an atom would acquire if it formed ions in aqueous solution (Hogness and Johnson, 1954, p. 90). This definition presents no difficulty in assigning valence numbers to atoms in ionically bonded compounds. However, molecules with predominantly covalent bonds do not dissociate in water and the valence numbers assigned to elements in covalently bonded compounds are therefore somewhat fictitious. We avoid these difficulties by adopting additional conventions regarding the assignment of valence numbers.

1. The valence number of all elements in pure form is equal to zero.

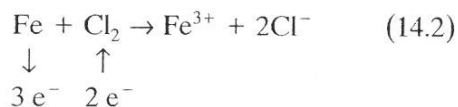
2. The valence number of H is +1, except in metal hydrides (e.g., LiH) in which it is -1.
3. The valence number of O is -2, except in peroxides (e.g., H₂O₂) in which it is -1.
4. Valence numbers are assigned to elements in molecules or complex ions in such a way that the algebraic sum of the valence numbers of the atoms of a neutral molecule is zero, or equals the charge of the complex ions.

The application of these rules is illustrated in Table 14.1. The examples listed there show that S can have valence numbers of +6, +4, 0, and -2. The valence numbers of C can be +4, 0, -2, and -4, those of Mn can be +7, +4, +3 and +2, and N can have valence numbers of +5 or -3. These values are derived from the chosen compounds or ions by the rules listed above, but do not necessarily represent all of the valence numbers these elements can assume.

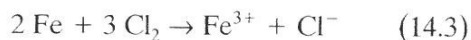
The first step in balancing an equation representing an oxidation-reduction reaction is to assign valence numbers to all of the elements that are involved. For example, Cl₂ gas reacts with metallic Fe in the presence of water to form Cl⁻ and Fe³⁺ ions:



The valence numbers of metallic Fe and Cl₂ are zero in accordance with the convention, whereas those of the ions are equal to their charges. Evidently, Fe gives up three electrons per atom to form Fe³⁺, whereas Cl₂ accepts two electrons per molecule to form 2 Cl⁻. The loss and gain of electrons is conveniently indicated by arrows:



where an arrow pointing down means *loss* and an arrow pointing up means *gain*. In order to balance the loss and gain of electrons, we cross multiply:

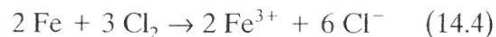


The Fe now gives up $2 \times 3 = 6$ electrons and the Cl₂ gains $3 \times 2 = 6$ electrons. Therefore, the elec-

Table 14.1 Assignment of Valence Numbers of Atoms in Chemical Compounds

Formula	Compound	Valence numbers
H ₂ SO ₄	sulfuric acid	H = +1 O = -2 S = +6
SO ₂	sulfur dioxide	O = -2 S = +4
H ₂ S	hydrogen sulfide	H = +1 S = -2
S ₂	sulfur gas	S = 0
CaCO ₃	calcium carbonate	O = -2 Ca = +2 C = +4
CO ₂	carbon dioxide	O = -2 C = +4
CO	carbon monoxide	O = -2 C = +2
CH ₄	methane	H = +1 C = -4
C	graphite	C = 0
FeCr ₂ O ₄	chromite	O = -2 Fe = +2 Cr = +3
MnO ₄ ⁻	permanganate ion	O = -2 Mn = +7
MnO ₂	pyrolusite	O = -2 Mn = +4
Mn ₃ O ₄	hausmannite	O = -2 Mn = +3 Mn = +2
HNO ₃	nitric acid	O = -2 H = +1 N = +5
NH ₃	ammonia	H = +1 N = -3
NH ₄ ⁺	ammonium	H = +1 N = -3

trons are in balance, and we can proceed with the mass balance of this equation:

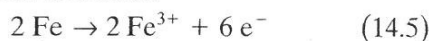


Note that the equation is electrically balanced as it must be. Equation 14.4 represents an oxidation-reduction reaction because both oxidation and reduction take place. The metallic Fe loses

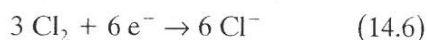
three electrons and is thereby *oxidized*, whereas the $\text{Cl}_2(\text{g})$ gains two electrons and is *reduced*. The electrons picked up by Cl_2 originate from the Fe, which acts as the *reducing agent*. The $\text{Cl}_2(\text{g})$ picks up the electrons given up by Fe and thereby acts as the *oxidizing agent*.

In general, metals are reducing agents (electron donors) and nonmetals are oxidizing agents (electron acceptors). However, the anions of nonmetals can also act as electron donors (reducing agents) because they may give up the electrons they acquired during their formation. Similarly, the cations of metals can be electron acceptors (oxidizing agents) to make up for the loss of electrons that occurred during their formation. These relationships are summarized in Table 14.2.

The oxidation-reduction reaction represented by equation 14.4 can be divided into two complementary half-reactions:



and

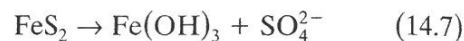


Equation 14.5 is the electron-donor (oxidation) half-reaction and equation 14.6 is the electron-acceptor (reduction) half-reaction. When the two half-reactions are added, the electrons cancel and we obtain equation 14.4 for the complete oxidation-reduction reaction.

Table 14.2 Schematic Representation of the Properties of Elements in Oxidation-Reduction Reactions

<i>Elements</i>	<i>Electron donors, reducing agents</i>	<i>Electron acceptors, oxidizing agents</i>
Metals	×	
Cations of metals		×
Nonmetals		×
Anions of nonmetals	×	
Nonmetals in complex anions (NO_3^- , SO_4^{2-} , PO_4^{3-})		×
Nonmetals in complex cations (NH_4^+)	×	

An important oxidation-reduction reaction in nature is the conversion of pyrite (FeS_2) into a precipitate of $\text{Fe}(\text{OH})_3$ and SO_4^{2-} ions in solution. In order to construct the equation for this reaction we treat FeS_2 as a reactant and $\text{Fe}(\text{OH})_3$ and SO_4^{2-} as products:



Equation 14.7 indicates that Fe experiences a change in its valence number from +2 in pyrite to +3 in ferric hydroxide. Therefore, the reaction liberates electrons that *must* be taken up by an electron acceptor (oxidizing agent). Natural environments may contain several different electron acceptors, which may be nonmetals like $\text{O}_2(\text{g})$ of the atmosphere, or cations such as Fe^{3+} , or complex ions containing oxidized forms of nonmetals like NO_3^- , SO_4^{2-} , or PO_4^{3-} . Some of these potential electron acceptors may not be "strong" enough to take an electron away from Fe^{2+} , but we leave that question for later consideration.

Diatomic O_2 molecules are the most common oxidizing agents in all natural environments in contact with the atmosphere. Therefore, we add O_2 as a reactant to equation 14.7:

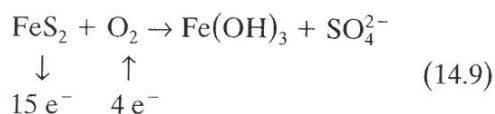


Next, we assign valence numbers and determine the number of electrons released and taken up in the reaction.

<i>Compound</i>	<i>Element</i>	<i>Valence number</i>
FeS_2	Fe	+2
	S	-1
O_2	O	0
$\text{Fe}(\text{OH})_3$	Fe	+3
	O	-2
	H	+1
SO_4^{2-}	S	+6
	O	-2

Note that the valence number of S in pyrite is -1, whereas in the SO_4^{2-} ion it is +6. Therefore, the S in the pyrite is an electron donor and provides

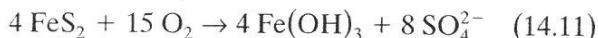
seven electrons per atom. The transformation of Fe^{2+} to Fe^{3+} provides an additional electron per atom of Fe. Therefore, the conversion of one mole of pyrite into $\text{Fe}(\text{OH})_3$ and SO_4^{2-} ions releases $(7 \times 2) + 1 = 15$ moles of electrons. The O_2 molecules accept four electrons per molecule because each O atom can take two electrons, which fill vacancies in its $2p$ orbital. The resulting electronic configuration of O^{2-} is $1s^2 2s^2 2p^6$, which is identical to that of Ne (Chapter 5). The electrons given off and accepted are indicated by means of arrows:



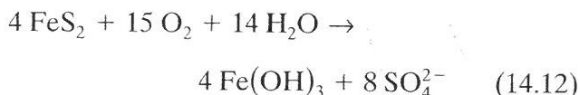
In order to balance the transfer of electrons, we cross multiply.



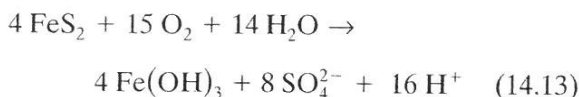
Next, we balance Fe and S:



In order to balance O, we add $14 \text{H}_2\text{O}$ to the left side:



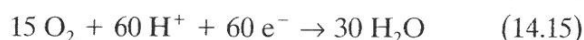
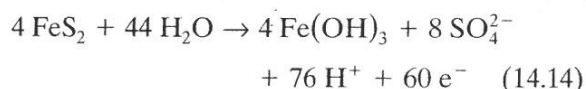
In order to balance H, we add 16H^+ to the right side:



Equation 14.13 indicates that the conversion of one mole of pyrite into solid ferric hydroxide and sulfate ion releases four moles of H^+ into the environment. The reaction occurs naturally when pyrite weathers by exposure to oxygenated water at or near the surface of the Earth. This process is greatly accelerated when pyrite-bearing rocks are exposed to weathering as a result of mining or quarrying. Pyrite occurs not only in sulfide ores associated with igneous and metamorphic rocks but also in shales that may be interbedded with coal seams, especially those that formed in marine coastal swamps. Stripmining of such coal seams

commonly causes acidification of local surface and groundwater because of the oxidation of pyrite (or its polymorph marcasite) by exposure to oxygenated water. The uncontrolled discharge of acid mine waters into streams or lakes is harmful to plant and animal life and may be accompanied by deposition of ferric hydroxide and even gypsum ($\text{CaSO}_4 \cdot 2\text{H}_2\text{O}$) where sufficient Ca^{2+} is present.

Even a complicated oxidation-reduction reaction like 14.13 can be divided into two complementary half-reactions representing oxidation and reduction, respectively:



The Fe^{2+} and S^{1-} of pyrite are the electron donors (reducing agents) and are themselves oxidized to Fe^{3+} and S^{6+} . The O_2 is the electron acceptor (oxidizing agent) and is reduced to O^{2-} .

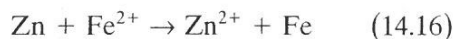
An oxidation-reduction reaction can, of course, achieve a state of chemical equilibrium, in which case it obeys the Law of Mass Action. The value of the equilibrium constant can be calculated from the standard free energy change of the reaction using data from Appendix B. However, in reality the oxidation of pyrite commonly does not reach equilibrium at or near the surface of the Earth because SO_4^{2-} and H^+ are lost from the system by the movement of water. Moreover, the amount of O_2 available to the reaction is commonly unlimited, allowing the reaction to continue until all of the available pyrite is consumed.

14.2 The Electromotive Series

In the preceding section we alluded to the fact that an oxidizing agent (electron acceptor) may not be strong enough to remove electrons from a particular reducing agent (electron donor). The ability of elements to act as electron donors or electron acceptors arises basically from the extent to which their orbitals are filled with electrons.

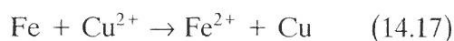
More specifically, this property depends on the decrease in the energy of the atoms that results from having only completely filled orbitals. Metallic elements empty partially filled orbitals, whereas nonmetals add electrons to orbitals that are nearly filled. For example, Na^+ and F^- both have the same electronic structure as Ne, but Na achieves this condition by losing an electron, whereas F must gain an electron.

Thermodynamics enables us to rank the elements in terms of their strengths as reducing agents or oxidizing agents. In order to maintain a consistent point of view in this presentation, we consider all elements to be *reducing agents* of varying strengths. Since metals have a natural tendency to be reducing agents, we consider how metals interact with each other. For example, suppose we place metallic Zn into a solution containing Fe^{2+} :

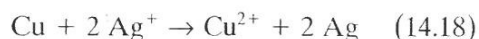


If the reaction proceeds as written, Zn atoms are able to force Fe^{2+} to accept two electrons. Consequently Zn atoms dissolve as Zn^{2+} and the Fe^{2+} ions precipitate as metallic Fe. However, the reaction could also run in the opposite direction, in which case metallic Fe could displace Zn^{2+} from the solution. We resolve the question by calculating $\Delta G_R^\circ = -16.29$ kcal for reaction 14.16 as written. Since free energy is released by this reaction, we conclude that in the standard state Zn is a stronger reducing agent than Fe because it is able to force Fe^{2+} to accept its two electrons.

Next, we consider the reaction:

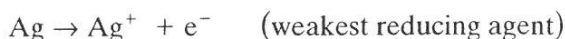
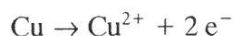
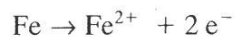
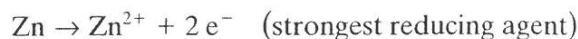


and find that $\Delta G_R^\circ = -34.51$ kcal. Evidently, Fe displaces Cu^{2+} ions from solution in the standard state. Therefore Fe is a stronger reducing agent than Cu. Finally, we add metallic Cu to a solution containing Ag^+ :



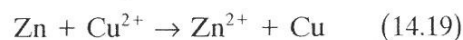
In this case, two Ag^+ ions are needed to take the two electrons that Cu atoms can give up. For this reaction, $\Delta G_R^\circ = -21.21$ kcal, which means that Cu is a stronger reducing agent than Ag.

The displacement reactions we have considered (14.16–14.18) permit us to rank the metals in order of decreasing strength as reducing agents in the so-called *electromotive series*:

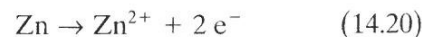


Zinc is the strongest reducing agent in this group of metals and displaces the ions of all metals located below it in the electromotive series. Silver is the weakest reducing agent in this group of elements, but it does displace ions of elements below it in the electromotive series. Conversely, because Ag atoms are the weakest reducing agent in the series of metals listed, Ag^+ is the strongest oxidizing agent among the ions we are considering.

The displacement reactions 14.16–14.18 are oxidation–reduction reactions involving the transfer of electrons from an electron donor to an electron acceptor and can be expressed as two complementary half-reactions. For example, when metallic Zn displaces Cu^{2+} from solution:



the reaction can be divided into two half-reactions:



These half-reactions can be separated from each other in an experimental setup called an *electrochemical cell*, which, in this case, consists of a Zn electrode and a Cu electrode dipping into an electrolyte solution containing Cu^{2+} ions. Figure 14.1 shows that the two electrodes are connected by a conducting wire through a voltmeter, which measures the electrical potential or voltage difference between the two electrodes. The voltage arises because electrons are produced at the Zn electrode according to equation 14.20, whereas electrons are consumed at the Cu electrode as shown by equation 14.21. The electrical potential generated by the half-reactions is called the *electromotive force*.

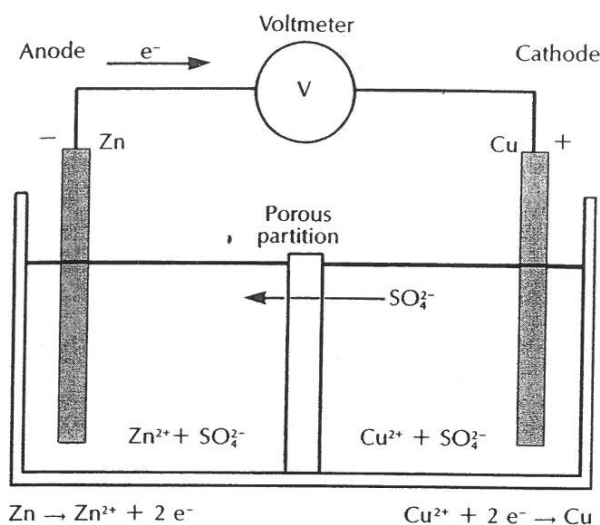


Figure 14.1 Schematic diagram of a Zn-Cu electrochemical cell. The two metal electrodes dip into solutions of ZnSO_4 and CuSO_4 , which are prevented from mixing by a porous partition. Metallic Zn gives up two electrons per atom and dissolves as Zn^{2+} . The electrons flow to the Cu electrode where Cu^{2+} ions pick up two electrons each and plate out as metallic Cu. Sulfate ions migrate through the porous partition to maintain the electrical neutrality of the solutions. As electrons flow from the Zn electrode to the Cu electrode, the activity of Zn^{2+} increases, whereas that of Cu^{2+} decreases until equilibrium is achieved at which time the flow of electrons stops.

(*emf*). The magnitude of the *emf* varies with time as the oxidation-reduction reaction composed of the two complementary electrode half-reactions approaches equilibrium. When the reaction is in the standard state ($[\text{Zn}^{2+}] = [\text{Cu}^{2+}] = 1.0$, $T = 25^\circ\text{C}$, $P = 1\text{ atm}$) the value of the *emf* is fixed by the definition of the standard state. When the reaction proceeds from the standard state toward equilibrium, the *emf* decreases and becomes equal to zero when equilibrium is achieved.

At equilibrium the Zn-Cu electrochemical cell obeys the Law of Mass Action and:

$$\frac{[\text{Zn}^{2+}]}{[\text{Cu}^{2+}]} = K \quad (14.22)$$

The magnitude of the equilibrium constant (K) can be calculated from the standard free energy

change of the reaction. In this case, $\Delta G_R^\circ = -50.8\text{ kcal}$ and $K = 10^{37.24}$. Evidently, when most of the Cu^{2+} ions have been driven out of the solution such that $[\text{Zn}^{2+}]/[\text{Cu}^{2+}] = 10^{37.24}$, the *emf* of the Zn-Cu cell becomes equal to zero. Note that at equilibrium the free energy change of the reaction is also equal to zero and that both ΔG_R and the *emf* change from fixed values in the standard state to zero at equilibrium.

The relationship between ΔG_R of an oxidation-reduction reaction in any state and the corresponding *emf* (E) is given by:

$$\Delta G_R = n\mathcal{F}E \quad (14.23)$$

where $\mathcal{F}$ is the Faraday constant and n is the number of electrons transferred in the reaction. The value of $\mathcal{F}$ is $96,489 \pm 20$ coulombs per mole or $23.06\text{ kcal per volt per gram equivalent}$. Since we can calculate the free energy change of an oxidation-reduction reaction in the standard state, we can use equation 14.23 to determine the corresponding *standard emf* (E°). For the Zn-Cu cell, $\Delta G_R^\circ = -50.8\text{ kcal}$ and therefore:

$$E^\circ = \frac{\Delta G_R^\circ}{n\mathcal{F}} = \frac{-50.8}{2 \times 23.06} = -1.10\text{ volts (V)} \quad (14.24)$$

The minus sign originates from the sign of ΔG_R° and contains information about the electrical polarity of the electrodes.

The relationship between ΔG_R° and E° can be used to rank the elements and their ions in terms of their strengths as reducing agents. In order to establish such a ranking, the *emf* generated by the dissociation of H_2 gas into H^+ in the standard state is arbitrarily set equal to zero:

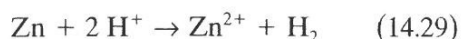
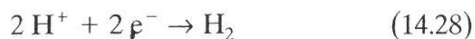
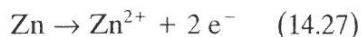


If $E^\circ = 0.00\text{ V}$, then ΔG_R° for this half-reaction must also be zero, which requires that we adopt the convention that the *standard* free energies of formation H^+ and e^- are equal to zero:

$$G_f^\circ(\text{H}^+) = G_f^\circ(\text{e}^-) = 0.00 \quad (14.26)$$

The standard hydrogen electrode can be connected to any metal electrode in the standard state to

form an electrochemical cell represented by an oxidation-reduction reaction that is obtained by adding the half-reactions. For example, if Zn is the metal electrode:



Equation 14.29 represents a displacement reaction and implies that Zn atoms can displace H^+ from solution and force it to form bubbles of H_2 gas. The value of ΔG_R° of equation 14.29 is calculated in the conventional manner:

$$\begin{aligned} \Delta G_R^\circ &= G_f^\circ(\text{Zn}^{2+}) + G_f^\circ(\text{H}_2) \\ &\quad - G_f^\circ(\text{Zn}) - 2G_f^\circ(\text{H}^+) \quad (14.30) \end{aligned}$$

However, $G_f^\circ(\text{H}_2)$ and $G_f^\circ(\text{H}^+)$ are both equal to zero by convention, so that ΔG_R° of reaction 14.29 arises *entirely* from the standard free energy change of the Zn electrode:

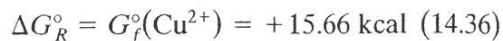
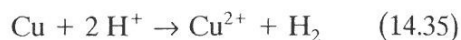
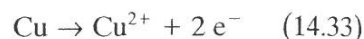
$$\begin{aligned} \Delta G_R^\circ &= G_f^\circ(\text{Zn}^{2+}) - G_f^\circ(\text{Zn}) \\ &= -35.14 - 0 = -35.14 \text{ kcal} \quad (14.31) \end{aligned}$$

Hence, the emf generated by the Zn- H_2 cell in the standard state is based *only* on the magnitude of the standard free energy change of the Zn electrode:

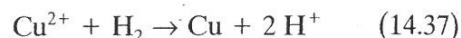
$$E^\circ = \frac{\Delta G_R^\circ}{n\mathcal{F}} = -\frac{35.14}{2 \times 23.06} = -0.76 \text{ V} \quad (14.32)$$

The emf generated by an electrode in the standard state relative to the H_2 electrode in the standard state is called the *standard electrode potential*. The standard electrode potential of the Zn electrode is -0.76 V , where the minus sign originates from the sign of ΔG_R° of the half-reaction. It indicates that when the Zn electrode is connected to the H_2 electrode in the standard state, the reaction proceeds in the direction indicated by the arrow in equation 14.29. In other words, Zn atoms force H^+ to accept its electrons; thus Zn is the *stronger reducing agent* and lies *above* H_2 in the electromotive series.

When the Cu electrode is connected to the H_2 electrode in the standard state:



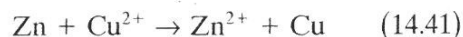
The positive sign indicates that reaction 14.35 actually proceeds in the opposite direction, that is, from right to left. Therefore, equation 14.35 should be reversed:



This means that H_2 is a stronger reducing agent than Cu because it forces Cu^{2+} from solution in the standard state. Therefore, the Cu electrode lies *below* the hydrogen electrode in the electromotive series. The standard electrode potential of the Cu electrode is calculated from the value of ΔG_R° of equation 14.35 by means of equation 14.23:

$$E^\circ = \frac{+15.66}{2 \times 23.06} = +0.34 \text{ V} \quad (14.38)$$

Note again that the H_2 electrode is "inert" and does not enter into the calculation. We can now recombine the Zn and Cu electrodes knowing that Zn is a stronger reducing agent than Cu:



$$\begin{aligned} E_{\text{Zn}+\text{Cu}}^\circ &= E_{\text{Zn}}^\circ + E_{\text{Cu}}^\circ = -0.76 + (-0.34) \\ &= -1.10 \text{ V} \quad (14.42) \end{aligned}$$

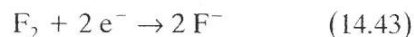
Note that we change the sign of E_{Cu}° when we inverted the Cu electrode half-reaction because it acts as the electron acceptor in this case.

The manipulation of electrode half-reactions, illustrated above, indicates that the electromotive series can be derived from standard free energy changes of the half-reactions, provided these are written with the *electrons on the right side*. When this convention is obeyed, the algebraic sign of

ΔG_R° indicates the position of the electrode above or below the H_2 electrode in the electromotive series. A minus sign means that the electrode is a *stronger* reducing agent than H_2 , whereas a positive sign indicates that it is *weaker* than the hydrogen electrode as a reducing agent. The numerical values of E° of the electrode half-reactions then permit us to rank elements such that the strongest reducing agent is at the top and the weakest reducing agent is at the bottom of the electromotive series.

The electromotive series in Table 14.3 is a listing of electrode half-reactions in order of decreasing strength of the reducing agents in the standard state. All of the reducing agents *above* the H_2 electrode are *stronger* than H_2 , and their half-reactions have negative standard electrode potentials. Those *below* the H_2 electrode are *weaker* reducing agents than H_2 . The algebraic sign and magnitude of the standard electrode potentials of the electrodes are derived from the sign and magnitude of ΔG_R° . Note that the alkali metals and alkaline earths are all strong reducing agents because these metals are good electron donors. Some members of the transition groups are moderate to weak reducing agents compared to H_2 (Zn, Cr, Fe, Ni, and Pb), whereas others are actually weaker than H_2 (Bi, Cu, Hg, Ag, and Pt).

The electrodes with positive standard electrode potentials act as *electron acceptors* in the standard state and therefore are *oxidizing agents*. For example, the standard electrode potential of the fluorine electrode in Table 14.3 is +2.88 V. The positive sign and large value indicate that F^- has only a weak tendency to give up electrons to form F_2 gas, whereas F_2 has a strong tendency to act as an electron acceptor or oxidizing agent:



Evidently, the *standard electrode potentials* reflect the same property of the elements as *electronegativities* (Section 7.2), and both arise from the tendencies of atoms to have completely filled orbitals by either giving up or accepting electrons.

Table 14.3 Electromotive Series Presented as Electrode Half-reactions in Order of Decreasing Strengths as Reducing Agents^a

Reducing agent	Oxidizing agent		Standard electrode potential, V
Cs	$\rightarrow Cs^+$	$+e^-$	-3.03
Li	$\rightarrow Li^+$	$+e^-$	-3.04
K	$\rightarrow K^+$	$+e^-$	-2.94
Ba	$\rightarrow Ba^{2+}$	$+2e^-$	-2.91
Sr	$\rightarrow Sr^{2+}$	$+2e^-$	-2.90
Ca	$\rightarrow Ca^{2+}$	$+2e^-$	-2.87
Na	$\rightarrow Na^+$	$+e^-$	-2.71
Rb	$\rightarrow Rb^+$	$+e^-$	-2.60
Y	$\rightarrow Y^{3+}$	$+3e^-$	-2.40
Mg	$\rightarrow Mg^{2+}$	$+2e^-$	-2.36
La	$\rightarrow La^{3+}$	$+3e^-$	-2.36
Ce	$\rightarrow Ce^{3+}$	$+3e^-$	-2.32
Sc	$\rightarrow Sc^{3+}$	$+3e^-$	-2.03
Be	$\rightarrow Be^{2+}$	$+2e^-$	-1.97
Th	$\rightarrow Th^{4+}$	$+4e^-$	-1.83
Al	$\rightarrow Al^{3+}$	$+3e^-$	-1.70
U	$\rightarrow U^{4+}$	$+4e^-$	-1.38
Mn	$\rightarrow Mn^{2+}$	$+2e^-$	-1.18
Nb	$\rightarrow Nb^{3+}$	$+3e^-$	-1.10
V	$\rightarrow V^{3+}$	$+3e^-$	-0.87
Zn	$\rightarrow Zn^{2+}$	$+2e^-$	-0.76
Cr	$\rightarrow Cr^{3+}$	$+3e^-$	-0.74
S^{2-}	$\rightarrow S$	$+2e^-$	-0.44
Fe	$\rightarrow Fe^{2+}$	$+2e^-$	-0.41
Cd	$\rightarrow Cd^{2+}$	$+2e^-$	-0.40
Co	$\rightarrow Co^{2+}$	$+2e^-$	-0.28
Ni	$\rightarrow Ni^{2+}$	$+2e^-$	-0.24
Mo	$\rightarrow Mo^{3+}$	$+3e^-$	-0.20
Sn	$\rightarrow Sn^{2+}$	$+2e^-$	-0.14
Pb	$\rightarrow Pb^{2+}$	$+2e^-$	-0.13
H_2	$\rightarrow 2H^+$	$+2e^-$	0.00
Bi	$\rightarrow Bi^{3+}$	$+3e^-$	+0.29
Cu	$\rightarrow Cu^{2+}$	$+2e^-$	+0.34
Cu	$\rightarrow Cu^+$	$+e^-$	+0.52
$2I^-$	$\rightarrow I_2$	$+2e^-$	+0.53
Se^{2-}	$\rightarrow Se$	$+2e^-$	+0.67
Ag	$\rightarrow Ag^+$	$+e^-$	+0.80
Hg	$\rightarrow Hg^{2+}$	$+e^-$	+0.85
Pd	$\rightarrow Pd^{2+}$	$+2e^-$	+0.92
$2Br^-$	$\rightarrow Br_2$	$+2e^-$	+1.08
Pt	$\rightarrow Pt^{2+}$	$+2e^-$	+1.19
$2Cl^-$	$\rightarrow Cl_2$	$+2e^-$	+1.36
Au	$\rightarrow Au^+$	$+e^-$	+1.69
Pt	$\rightarrow Pt^+$	$+e^-$	+2.64
$2F^-$	$\rightarrow F_2$	$+2e^-$	+2.88

^aBased on G_f° values in Appendix B.

14.3 The Emf of Electrochemical Cells

The relationship between the free energy change of a chemical reaction and the activities of the reactants and products is given by an equation derived in Section 11.8 (equation 11.73):

$$\Delta G_R - \Delta G_R^\circ = RT \ln Q \quad (14.44)$$

where ΔG_R is the free energy change of the reaction in any state, ΔG_R° is the free energy change in the standard state, R is the gas constant, T is the standard temperature in kelvins, and Q is the reaction quotient. We now convert ΔG values to emf by substituting equation 14.23 into equation 14.44. After rearranging terms and converting to logarithms to the base 10, we have:

$$E = E^\circ + \frac{2.303RT}{n\mathcal{F}} \log Q \quad (14.45)$$

Since $R = 1.987 \times 10^{-3}$ kcal/deg·mol, $T = 298.15$ K, $\mathcal{F} = 23.06$ kcal/V · number of electrons transferred:

$$E = E^\circ + \frac{2.303 \times (1.987 \times 10^{-3}) \times 298.15}{n \times 23.06} \log Q \quad (14.46)$$

$$E = E^\circ + \frac{0.05916}{n} \log Q \quad (14.47)$$

Equation 14.47 provides a relationship between the activities of the reactants and products of an oxidation-reduction reaction, expressed by Q , and the emf that would be generated by the reaction if it were split into two half-reactions in an electrochemical cell. When the reaction achieves equilibrium, $\Delta G_R = 0$, $E = 0$, and $Q = K$ where K is the equilibrium constant. Therefore, at equilibrium:

$$E^\circ = -\frac{0.05916}{n} \log K \quad (14.48)$$

Equation 14.48 relates the emf of an oxidation-reduction reaction in the standard state to the equilibrium constant of that reaction.

We can apply these relationships to the Zn-Cu electrochemical cell represented by equation 14.41. When this reaction is neither in the standard state nor at equilibrium, the emf it generates is given by:

$$E = E^\circ + \frac{0.05916}{2} \log \frac{[\text{Zn}^{2+}]}{[\text{Cu}^{2+}]} \quad (14.49)$$

where E° is the emf generated by the Zn-Cu cell in the standard state, which has a value of -1.10 V as calculated in equation 14.24. Therefore, when the reaction is proceeding toward equilibrium, the emf is:

$$E = -1.10 + 0.02958 \log \frac{[\text{Zn}^{2+}]}{[\text{Cu}^{2+}]} \quad (14.50)$$

In the standard state, $[\text{Zn}^{2+}] = [\text{Cu}^{2+}] = 1.0$ and therefore $\log [\text{Zn}^{2+}]/[\text{Cu}^{2+}] = 0$. Consequently, in the standard state the emf is:

$$E = E^\circ = -1.10 \text{ V} \quad (14.51)$$

At equilibrium, $[\text{Zn}^{2+}]/[\text{Cu}^{2+}] = 10^{37.24}$ (equation 14.22). Therefore, at equilibrium:

$$E = -1.10 + 0.02958 \times 37.24 = 0.00 \text{ V} \quad (14.52)$$

The progress of the reaction in the Zn-Cu cell is illustrated by plotting equation 14.50 in Figure 14.2 in coordinates of E and $\log [\text{Zn}^{2+}]/[\text{Cu}^{2+}]$. The straight line indicates how E changes from -1.10 V in the standard state ($\log [\text{Zn}^{2+}]/[\text{Cu}^{2+}] = 0$) to 0.00 V at equilibrium ($\log [\text{Zn}^{2+}]/[\text{Cu}^{2+}] = 37.24$). When an external voltage is applied to the Zn-Cu cell, the $[\text{Zn}^{2+}]/[\text{Cu}^{2+}]$ ratio changes until it is compatible with the applied voltage.

These concepts seem fairly removed from oxidation-reduction reactions in natural environments. However, they enable us to define a special kind of emf known as the E_h . The E_h is the emf generated between an electrode in any state and the H_2 electrode in the standard state. Since the H_2 electrode is in the standard state, it does not affect the magnitude of the E_h and acts only as a reference. We illustrate the usefulness of this concept by considering an electrode at which Fe^{2+} is oxidized to Fe^{3+} :

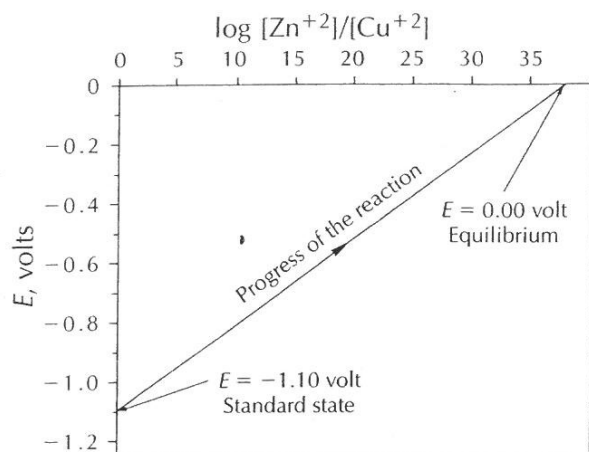
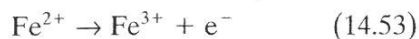
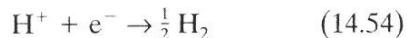


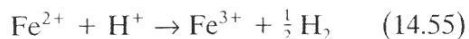
Figure 14.2 Variation of the emf of a Zn-Cu electrochemical cell as the oxidation-reduction reaction on which it is based progresses from the standard state when $[\text{Zn}^{2+}] = [\text{Cu}^{2+}] = 1.0 \text{ mol/L}$, and $E^\circ = -1.10 \text{ V}$ toward equilibrium when $[\text{Zn}^{2+}]/[\text{Cu}^{2+}] = 10^{37.5}$ and $E = 0.00 \text{ V}$.



The complete oxidation-reduction reaction, when this electrode is combined with the H_2 electrode in the standard state, is obtained by adding the H_2 half-reaction:



to equation 14.53, which yields:



Since the H_2 electrode is in the standard state, $[\text{H}^+] = [\text{H}_2] = 1.0$. Therefore, the reaction quotient of reaction 14.55 is:

$$Q = \frac{[\text{Fe}^{3+}]}{[\text{Fe}^{2+}]} \quad (14.56)$$

The Eh generated by the Fe^{2+} electrode is given by equation 14.47. When both electrodes are in the standard state, the change in free energy of equation 14.55 is

$$\begin{aligned} \Delta G_R^\circ &= [G_f^\circ(\text{Fe}^{3+}) + \frac{1}{2}G_f^\circ(\text{H}_2)] \\ &\quad - [G_f^\circ(\text{H}^+) + G_f^\circ(\text{Fe}^{2+})] \quad (14.57) \end{aligned}$$

Since $G_f^\circ(\text{H}_2)$ and $G_f^\circ(\text{H}^+)$ are both equal to zero, ΔG_R° of equation 14.55 depends entirely on the Fe^{2+} electrode. Therefore, we can calculate it *directly* from the electrode half-reaction. In this case, $\Delta G_R^\circ = +17.75 \text{ kcal}$ and the standard electrode potential $E^\circ = +0.769 \text{ V}$. The Eh generated by the Fe^{2+} electrode is related to the ratio of the activities of Fe^{3+} and Fe^{2+} by the equation:

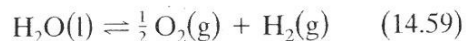
$$Eh = 0.769 + \frac{0.05916}{1} \log \frac{[\text{Fe}^{3+}]}{[\text{Fe}^{2+}]} \quad (14.58)$$

The equation establishes the relationship between the emf generated by the Fe^{2+} electrode connected to the H_2 electrode in the standard state for different values of the $[\text{Fe}^{3+}]/[\text{Fe}^{2+}]$ activity ratio.

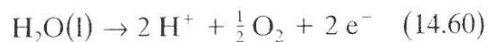
In natural systems the Eh , like pH , is an *environmental parameter* whose value reflects the ability of the natural system to be an electron donor or acceptor relative to the standard H_2 electrode. Therefore, equation 14.58 leads to the important conclusion that the $[\text{Fe}^{3+}]/[\text{Fe}^{2+}]$ ratio of a natural geochemical system depends on the Eh of the environment. This example suggests that the concepts of electrochemistry presented in Section 14.2 are indeed applicable to natural systems.

14.4 Stability Limits of Water in Terms of Eh and pH

We showed in Section 12.5 that liquid water must maintain an equilibrium with O_2 and H_2 gas:



We used this equation to set limits on the fugacity of O_2 required for stability of liquid water on the surface of the Earth. Now we want to determine the stability limits of water in terms of Eh and pH . At first glance, this may appear to be a pretty hopeless proposition until we recall that we are dealing here with electrode half-reactions. Therefore, we invent the *water electrode*:



In this electrode the O in the water molecule gives up two electrons and thereby changes its

valence number from -2 to 0 . Equation 14.60 therefore is an electron-donor half-reaction that can generate an Eh when it is connected to the hydrogen reference electrode, which is maintained in the standard state. It does not matter whether an $\text{H}_2\text{O}-\text{H}_2$ electrochemical cell can actually be constructed. Moreover, we do not need to complete the oxidation-reduction reaction by adding the standard H_2 electrode reaction to equation 14.60 because the Eh is generated entirely by the water electrode. Therefore, we can express the Eh of the water electrode by the equation:

$$\text{Eh} = E^\circ + \frac{0.05916}{2} \log([\text{O}_2]^{1/2}[\text{H}^+]^2) \quad (14.61)$$

We determine E° by calculating ΔG_R° for the water electrode as written in equation 14.60:

$$\begin{aligned} \Delta G_R^\circ &= [2G_f^\circ(\text{H}^+) + \frac{1}{2}G_f^\circ(\text{O}_2) + 2G_f^\circ(\text{e}^-)] \\ &\quad - [G_f^\circ(\text{H}_2\text{O}(\text{l}))] \quad (14.62) \end{aligned}$$

Since

$$G_f^\circ(\text{H}^+) = G_f^\circ(\text{O}_2) = G_f^\circ(\text{e}^-) = 0,$$

$$\Delta G_R^\circ = -G_f^\circ(\text{H}_2\text{O}(\text{l})) = +56.687 \text{ kcal} \quad (14.63)$$

and hence:

$$E^\circ = +\frac{56.687}{2 \times 23.06} = +1.23 \text{ V} \quad (14.64)$$

Therefore, the Eh generated by the water electrode is:

$$\text{Eh} = 1.23 + \frac{0.05916}{2} \log([\text{O}_2]^{1/2}[\text{H}^+]^2) \quad (14.65)$$

$$\begin{aligned} \text{Eh} &= 1.23 + \frac{0.05916}{2} \left(\frac{1}{2}\right) \log[\text{O}_2] \\ &\quad + \frac{0.05916 \times (2)}{2} \log[\text{H}^+] \quad (14.66) \end{aligned}$$

$$\text{Eh} = 1.23 + 0.01479 \log[\text{O}_2] - 0.05916 \text{ pH} \quad (14.67)$$

Equation 14.67 is very important because it establishes the relationship between the Eh and pH of a geochemical environment and the corresponding fugacity of O_2 .

We can also use equation 14.67 to state the stability limits of liquid water on the surface of the Earth. According to the discussion in Section 12.5, the fugacity of O_2 in equilibrium with liquid water can vary between 1.0 and $10^{-83.1}$ atm. When we apply these limits to equation 14.67, we obtain for $\text{O}_2 = 1.0$ atm:

$$\text{Eh} = 1.23 - 0.05916 \text{ pH} \quad (14.68)$$

and for $\text{O}_2 = 10^{-83.1}$ atm:

$$\text{Eh} = 1.23 + 0.01479(-83.1) - 0.05916 \text{ pH} \quad (14.69)$$

$$\text{Eh} = -0.05916 \text{ pH} \quad (14.70)$$

Equations 14.68 and 14.70 are straight lines in coordinates of Eh and pH, which define the stability field of liquid water in Figure 14.3. All natural chemical reactions that take place in the presence of water at 25°C and 1 atm pressure are restricted to that region of the diagram.

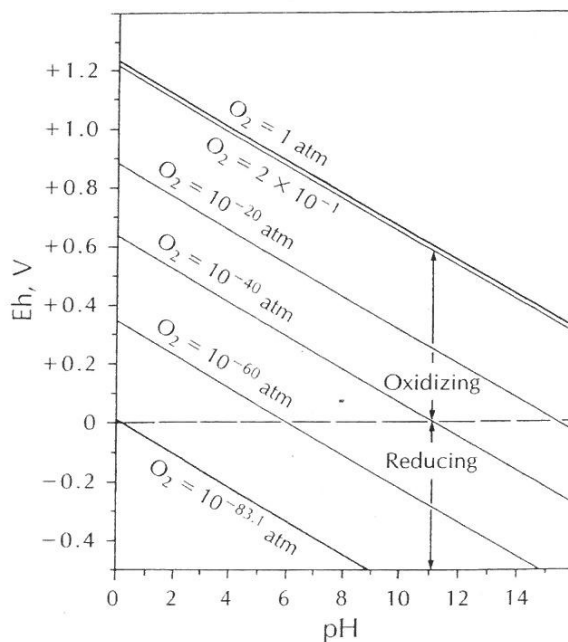


Figure 14.3 Stability field of liquid water at the surface of the Earth at 25°C and 1 atm pressure in coordinates of Eh and pH. The contours of O_2 fugacity are based on equation 14.67 and arise from the dissociation of liquid water into O_2 , H^+ , and electrons.

The stability field of liquid water in Figure 14.3 contains environments having positive as well as negative Eh values. A positive Eh implies that the environment is a *weaker* reducing agent than the standard H₂ electrode. Therefore, such environments actually function as *electron acceptors* (oxidizing agents) relative to the standard H₂ electrode. Hence, such environments are said to be *oxidizing environments*, meaning that they have the ability to *accept* electrons from electron donors.

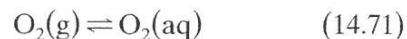
Environments having negative Eh values are *stronger* reducing agents than the standard H₂ electrode. Therefore, such environments function as *electron donors* (reducing agents) and are said to be *reducing environments*.

An oxidizing environment contains elements or compounds that are capable of accepting electrons in case an electron donor becomes exposed to it. The most common oxidizing agent at or near the surface of the Earth is atmospheric molecular oxygen, which dissolves in water and may be transported by the movement of water to environments that are not directly exposed to the atmosphere. In addition, the equilibrium of water with O₂ and H₂ requires the presence of both gases wherever water is present on the Earth. Consequently, some O₂ is present even in environments described as *anoxic*. However, the fugacity of O₂ in anoxic environments may be much less than is required to sustain life, although certain kinds of bacteria thrive in such environments.

The relationship of Eh, pH, and O₂, expressed by equation 14.67, enables us to calculate the fugacity of O₂ from the Eh and pH of an environment. Equation 14.67 can also be used to derive a series of contours for selected values of the O₂ fugacity, as shown in Figure 14.3. These contours reflect the fact that the O₂ fugacity of natural environments can vary over an extremely wide range from 1 to 10^{-83.1} atm. Even environments classified as *oxidizing* in Figure 14.3 may contain only very small amounts of O₂.

Consider an environment with pH = 7.0 and Eh = +0.6 V. Equation 14.67 indicates that in this environment [O₂] = 10^{-14.6} atm. The gaseous

O₂ is in equilibrium with O₂ molecules dissolved in water:

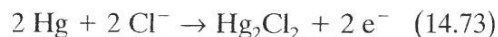


At equilibrium 25°C (Stumm and Morgan, 1996, p. 215):

$$\frac{[\text{O}_2(\text{aq})]}{[\text{O}_2(\text{g})]} = 10^{-2.9} \quad (14.72)$$

Therefore, [O₂(aq)] = 10^{-17.5} mol/L, which is equivalent to about 1 × 10⁻¹³ μg/mL of dissolved O₂. Most fish require more than one million times more O₂. Nevertheless, the environment is classified as being *oxidizing*.

The Eh of an environment can be measured directly as the voltage between a shiny Pt electrode inserted into the solution to be tested and a reference electrode. The reference electrode used most frequently is the *calomel electrode* in which Hg reacts with Cl⁻ to form Hg₂Cl₂ plus electrons:



The emf of this electrode is stabilized by maintaining the activity of Cl⁻ at a constant value by immersing it in a saturated solution of KCl. However, in practice, Eh measurements are difficult to make because the insertion of the electrode may disturb the environment, or because the environment is inaccessible, and for other reasons discussed by Garrels and Christ (1965) and by Stumm and Morgan (1996). Nevertheless, sufficient Eh measurements exist to characterize common geological environments in terms of the range of Eh and pH values (Baas-Becking et al., 1960). A schematic representation of these data in Figure 14.4 indicates that environments in contact with the atmosphere (rainwater, streams, oceans, and mine waters) have positive Eh values, whereas environments that are isolated from the atmosphere (waterlogged soils, euxenic marine basins, and organic-rich brines) have negative Eh values. However, many environments are transitional between these extremes and have intermediate Eh values.

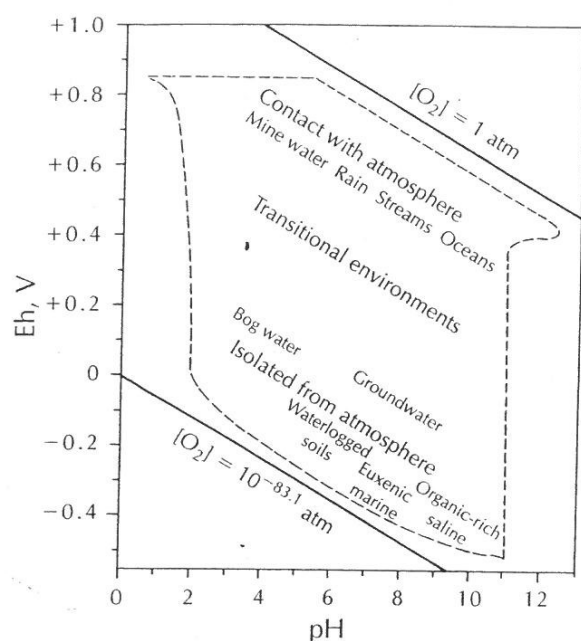


Figure 14.4 Range of Eh and pH conditions in natural environments based on data by Baas-Becking et al. (1960) and modified after Garrels and Christ (1965).

14.5 Stability of Iron Compounds

The stability of solid compounds of elements having different oxidation states can be represented in terms of Eh and pH when they participate in electrode half-reactions. We therefore demonstrate the construction of such *Eh-pH stability diagrams* by using the naturally occurring compounds of Fe as an example. Such diagrams were originally adapted by Garrels (1960) from electrochemical theory used in the study of the corrosion of metals (Pourbaix, 1949). Eh-pH diagrams are widely used in geochemistry to study the formation and weathering of ore deposits (Garrels and Christ, 1965), environmental geochemistry, and nuclear waste disposal (Brookins, 1978). A book devoted exclusively to this subject was published by Brookins (1988).

a. Oxides of Iron

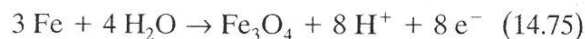
The naturally occurring oxides of Fe include magnetite (Fe_3O_4) and hematite ($\alpha\text{-Fe}_2\text{O}_3$). Other oxides such as maghemite ($\gamma\text{-Fe}_2\text{O}_3$) and oxyhy-

droxides goethite ($\alpha\text{-FeOOH}$) and lepidocrocite ($\gamma\text{-FeOOH}$) occur metastably and, in time, recrystallize to hematite. Nevertheless, their stability in coordinates of Eh and pH can be delineated by the same methods that apply to hematite and magnetite.

We begin the presentation with the oxidation of metallic Fe to magnetite, which we express as an electrode half-reaction:



Note that magnetite is a mixed oxide composed of FeO and Fe_2O_3 . Therefore, one of the metallic Fe atoms loses two electrons to form FeO while the other two atoms lose three electrons each to form Fe_2O_3 . Therefore, the three moles of Fe required to make one mole of magnetite give up a total of eight moles of electrons. In order to account for the O of magnetite we add four moles of H_2O to the left side of the equation, which requires the appearance of eight moles of H^+ on the right side:



This electrode generates an emf when it is connected to the standard H_2 electrode:

$$\text{Eh} = E^\circ + \frac{0.05916}{8} \log [\text{H}^+]^8 \quad (14.76)$$

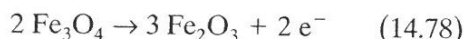
Since ΔG_R° for this electrode is -15.85 kcal, $E^\circ = -0.086$ V. By substituting this value into equation 14.76 and after converting to pH, we obtain:

$$\text{Eh} = -0.086 - 0.0592 \text{pH} \quad (14.77)$$

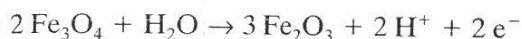
The straight line represented by this equation relates the Eh of the Fe- Fe_3O_4 electrode to the activity of H^+ . If the Eh is fixed by the environment, $[\text{H}^+]$ is uniquely determined by equation (14.77) in terms of the corresponding pH value. If the $[\text{H}^+]$ of the environment is greater than required by equation 14.77, the electrode half-reaction reduces magnetite to metallic Fe in accordance with equation 14.75. Therefore, metallic Fe is the stable phase under conditions represented by all points located left of the

Eh-pH line for the Fe-Fe₃O₄ electrode. If [H⁺] of the environment having a fixed Eh value is less than that required by the electrode (equation 14.75), metallic Fe is converted to Fe₃O₄. Therefore, magnetite is the stable phase in environments represented by points located to the right of the Eh-pH line for the Fe-Fe₃O₄ electrode. However, metallic Fe is not stable on the surface of the earth because its stability field lies *outside* the stability limits of water. We reached the same conclusion during the construction of the fugacity diagram for Fe oxides in Section 12.5b.

The conversion of magnetite to hematite can be expressed as an electrode half-reaction:



We require one mole of H₂O on the left side to balance O and then obtain 2 H⁺ on the right:



$$\Delta G_R^\circ = +9.087 \text{ kcal} \quad E^\circ = +0.20 \text{ V} \quad (14.79)$$

Therefore, the Eh generated by this electrode as a function of pH is:

$$\text{Eh} = 0.20 + \frac{0.05916}{2} \log [\text{H}^+]^2 \quad (14.80)$$

which reduces to:

$$\text{Eh} = 0.20 - 0.0592 \text{ pH} \quad (14.81)$$

Equations 14.77 and 14.81 define the stability limits of metallic Fe, magnetite, and hematite in Figure 14.5. Magnetite is stable and can precipitate from aqueous solution primarily in reducing environments having Eh < 0.0 V. Hematite is stable under oxidizing conditions, but its stability field does extend into reducing environments having pH > 4.6.

b. Solubility of Iron Oxides

Both magnetite and hematite are slightly soluble in water depending on the pH (Section 12.5b, Figure 12.7). Several different ionic and molecular species can form, but Fe²⁺ and Fe³⁺ dominate at low pH. The solubility of Fe-oxides limits their

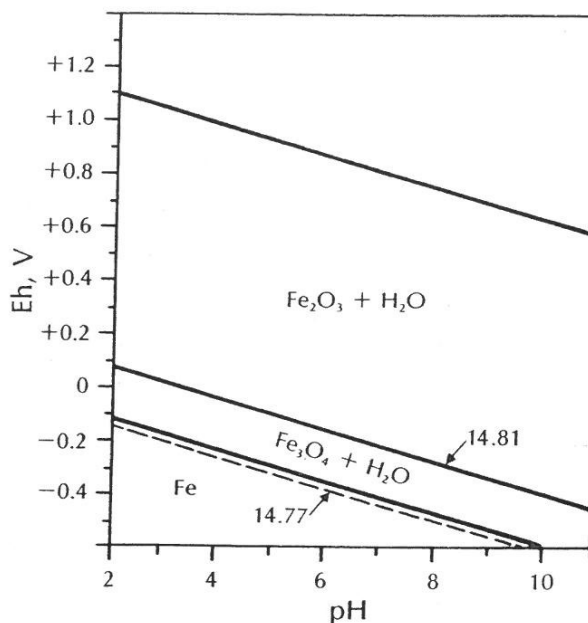
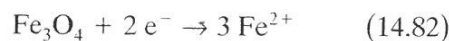


Figure 14.5 Stability limits of metallic Fe, magnetite, and hematite in the presence of water at 25°C in coordinates of Eh and pH. Note that the pH scale starts at 2 rather than 0.

stability in natural environments because they may be removed by the movement of acidic solutions in the course of geologic time.

We consider first the solubility of magnetite and hematite with respect to Fe²⁺. The solution of magnetite to form Fe²⁺ involves a change in the valence number of the trivalent Fe and therefore can be expressed as an electrode half-reaction:



We must reverse this equation in order to keep the electrons on the right side and, by adding H₂O and H⁺ as required, we obtain:



$$\Delta G_R^\circ = +40.698 \text{ kcal} \quad E^\circ = +0.88 \text{ V} \quad (14.83)$$

Therefore,

$$\text{Eh} = 0.88 + \frac{0.05916}{2} \log \left(\frac{[\text{H}^+]^8}{[\text{Fe}^{2+}]^3} \right) \quad (14.84)$$

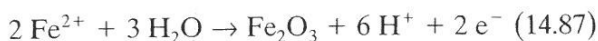
which reduces to:

$$\begin{aligned} \text{Eh} = 0.88 + \frac{0.05916 \times 8}{2} \log [\text{H}^+] \\ - \frac{0.05916 \times 3}{2} \log [\text{Fe}^{2+}] \end{aligned} \quad (14.85)$$

$$\text{Eh} = 0.88 - 0.237 \text{ pH} - 0.089 \log [\text{Fe}^{2+}] \quad (14.86)$$

Equation 14.86 represents a family of parallel lines in coordinates of Eh and pH corresponding to selected values of $[\text{Fe}^{2+}]$. These lines express the solubility of *magnetite* in the Eh-pH plane and therefore are valid only within the stability field of magnetite.

The solubility of *hematite* with respect to Fe^{2+} involves a change in the valence number of Fe from +3 to +2 and therefore can be written as an electrode half-reaction with the electrons on the right side:



$$\Delta G_R^\circ = +30.161 \text{ kcal} \quad E^\circ = +0.65 \text{ V}$$

$$\text{Eh} = 0.65 + \frac{0.05916}{2} \log \left(\frac{[\text{H}^+]^6}{[\text{Fe}^{2+}]^2} \right) \quad (14.88)$$

$$\text{Eh} = 0.65 - 0.177 \text{ pH} - 0.0592 \log [\text{Fe}^{2+}] \quad (14.89)$$

Equations 14.86 and 14.89 have been used to draw contours across the stability fields of magnetite and hematite in Figure 14.6 at activities of $\text{Fe}^{2+} = 10^{-4}$, 10^{-6} , and 10^{-8} mol/L. Evidently, the solubility of both oxides increases with decreasing environmental pH and Eh. Therefore, when hematite and magnetite are exposed to acidic environments with low Eh values, equivalent to low O_2 fugacities, they become *unstable* and go into solution.

Conversely, when a solution containing Fe^{2+} is neutralized (pH increases), or aerated (Eh increases), or both, hematite or magnetite may precipitate. For example, a solution containing $[\text{Fe}^{2+}] = 10^{-4}$ mol/L ($5.6 \mu\text{g/mL}$), whose Eh and pH are the coordinates of point P in Figure 14.6, is undersaturated with respect to hematite. If the environmental Eh and pH change along the

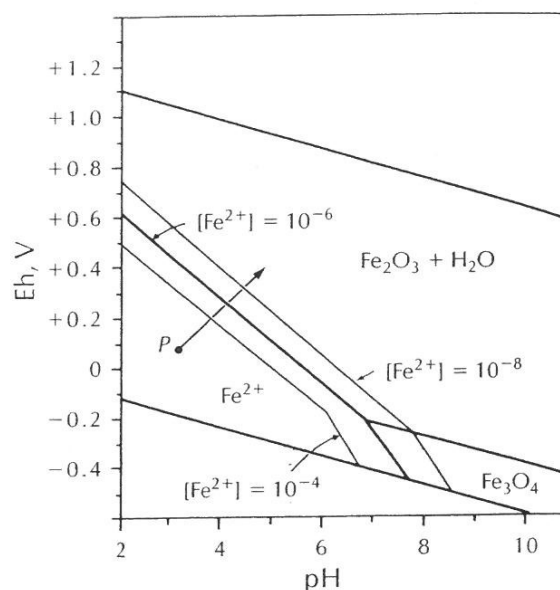


Figure 14.6 Solubility of hematite and magnetite with respect to Fe^{2+} represented by equations 14.89 and 14.86, respectively. An activity of 10^{-6} mol/L of ions is regarded as an effective limit on the stability of a solid phase because when $[\text{Fe}^{2+}] > 10^{-6}$ mol/L, the minerals can be removed from rocks by the movement of groundwater in the course of geologic time.

arrow in Figure 14.6, the solution becomes saturated with respect to hematite when the path reaches the 10^{-4} contour. If Eh and pH continue to change in the direction of the arrow, hematite precipitates continuously as the activity of Fe^{2+} decreases to 10^{-6} mol/L ($0.056 \mu\text{g/mL}$) and 10^{-8} mol/L ($0.00056 \mu\text{g/mL}$). The amount of hematite that precipitates during this process is 0.44 g/L as $[\text{Fe}^{2+}]$ in the solution decreases from 10^{-4} to 10^{-6} mol/L. If the environmental conditions continue to change in the direction of the arrow, the activity of Fe^{2+} in the solution decreases from 10^{-6} to 10^{-8} mol/L, but the amount of hematite that precipitates is only 0.0044 g/L. These figures justify the convention that an activity of 10^{-6} mol/L of ionic species is an effective limit on the stability of solids.

The solubility of hematite with respect to Fe^{3+} does not require a change in the valence

number of Fe. Therefore, we can construct the desired equation purely on the basis of mass balance:



$$\Delta G_R^\circ = +5.339 \text{ kcal} \quad K = 10^{-3.91}$$

$$\frac{[\text{Fe}^{3+}]^2}{[\text{H}^+]^6} = 10^{-3.91} \quad (14.91)$$

$$2 \log [\text{Fe}^{3+}] - 6 \log [\text{H}^+] = -3.91 \quad (14.92)$$

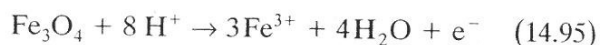
$$\log [\text{Fe}^{3+}] + 3 \text{pH} = -1.96 \quad (14.93)$$

If

$$\begin{aligned} [\text{Fe}^{3+}] &= 10^{-6} \text{ mol/L}, \\ \text{pH} &= 1.35 \end{aligned} \quad (14.94)$$

Evidently, the solubility of hematite with respect to Fe^{3+} is independent of Eh because it does not require an oxidation or reduction.

The solution of magnetite to form Fe^{3+} does require the oxidation of one divalent Fe^{2+} to the trivalent state and therefore must be expressed as an electrode half-reaction:



$$\Delta G_R^\circ = +12.552 \text{ kcal} \quad E^\circ = +0.54 \text{ V}$$

$$\text{Eh} = 0.54 + \frac{0.05916}{1} \log \left(\frac{[\text{Fe}^{3+}]^3}{[\text{H}^+]^8} \right) \quad (14.96)$$

$$\begin{aligned} \text{Eh} &= 0.54 + \frac{0.05916 \times 3}{1} \log [\text{Fe}^{3+}] \\ &\quad - \frac{0.05916}{1} \times 8 \log [\text{H}^+] \end{aligned} \quad (14.97)$$

$$\text{Eh} = 0.54 + 0.177 \log [\text{Fe}^{3+}] + 0.473 \text{pH} \quad (14.98)$$

$$\text{If } [\text{Fe}^{3+}] = 10^{-6} \text{ mol/L},$$

$$\text{Eh} = -0.52 + 0.473 \text{pH} \quad (14.99)$$

Equations 14.94 and 14.99 have been plotted in Figure 14.7A by superimposing them on Figure 14.6. The results reveal a discrepancy.

Equation 14.94 ($\log [\text{Fe}^{3+}] = 10^{-6} \text{ mol/L}$) cuts off a corner of the hematite field in the upper left corner of the diagram and intersects the

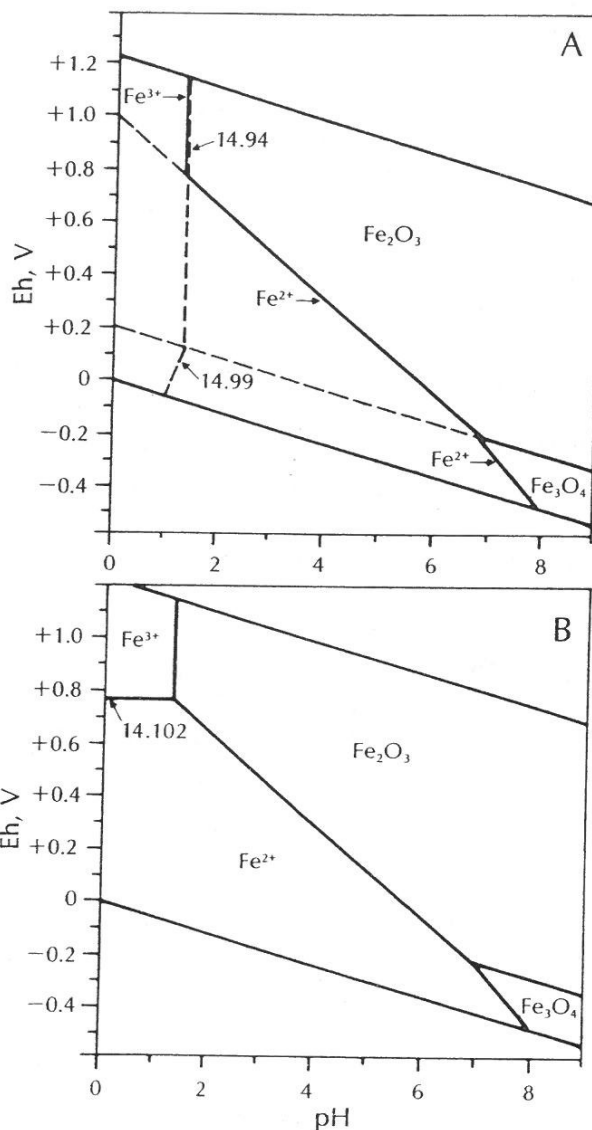


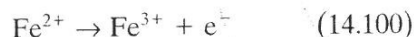
Figure 14.7 A: Solubility of hematite and magnetite at 25°C with respect to Fe^{3+} based on equations 14.94 and 14.98 drawn at $[\text{Fe}^{3+}] = 10^{-6} \text{ mol/L}$. The results indicate that $[\text{Fe}^{2+}] \gg [\text{Fe}^{3+}]$, except under acidic and strongly oxidizing conditions in the stability field of hematite.

B: Completed Eh-pH diagram for the oxides of Fe and their dominant ions at an activity of 10^{-6} mol/L and 25°C. Equation 14.102 fixes the boundary where $[\text{Fe}^{3+}] = [\text{Fe}^{2+}]$ in equilibrium with hematite. Note that the pH scale has been shifted compared to Figure 14.6.

$[\text{Fe}^{2+}] = 10^{-6}$ mol/L solubility contour at a point whose coordinates are $\text{pH} = 1.35$, $E_h = +0.77$ V. At the point of intersection, the activities of Fe^{3+} and Fe^{2+} are both equal to 10^{-6} mol/L. However, at lower E_h values the line (equation 14.94) enters the region of the diagram where $[\text{Fe}^{2+}] \gg 10^{-6}$ mol/L, whereas, $[\text{Fe}^{3+}]$ remains at 10^{-6} mol/L. Therefore, the solubility of hematite with respect to Fe^{2+} is much greater than its solubility to form Fe^{3+} in all environments where $E_h < +0.77$ V.

The solubility of magnetite in its own stability field also favors Fe^{2+} over Fe^{3+} . Equation 14.99 (for $\log [\text{Fe}^{3+}] = 10^{-6}$ mol/L) plots far to the left of the Fe^{2+} solubility boundary at 10^{-6} mol/L. Therefore, Fe^{3+} is much less abundant than Fe^{2+} in equilibrium with *hematite*, except if $\text{pH} < 1.35$ and $E_h > 0.77$ V, and it does *not* become dominant over Fe^{2+} in equilibrium with *magnetite* under any E_h - pH conditions.

The boundary along which $[\text{Fe}^{3+}]$ is equal to $[\text{Fe}^{2+}]$ in equilibrium with hematite can be drawn graphically by plotting contours at different activities and by connecting the points of intersection. A more elegant method is to write the electrode half-reaction (14.53):



Since $\Delta G_R^\circ = +17.75$ kcal and $E^\circ = 0.77$ V,

$$E_h = 0.77 + \frac{0.05916}{1} \log \left(\frac{[\text{Fe}^{3+}]}{[\text{Fe}^{2+}]} \right) \quad (14.101)$$

We are interested in environments where $[\text{Fe}^{3+}] = [\text{Fe}^{2+}]$, and $\log([\text{Fe}^{3+}]/[\text{Fe}^{2+}]) = 0$. Therefore, the boundary along which this condition occurs is given by:

$$E_h = 0.77 \text{ V} \quad (14.102)$$

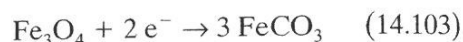
We see in Figure 14.7B that Fe^{3+} is the dominant ion under narrowly defined conditions ($\text{pH} < 1.35$ and $E_h > 0.77$ V) encountered only in some acid mine-waters.

c. Stability of Ferrous Carbonate

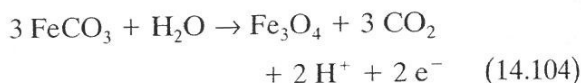
The dominance of Fe^{2+} over Fe^{3+} in most environments on the surface of the Earth is reflected

in the fact that *siderite* (FeCO_3) rather than $\text{Fe}_2(\text{CO}_3)_3$ is the common carbonate of Fe. Siderite may precipitate from saturated solutions which are in equilibrium with CO_2 gas or contain a fixed amount of carbonate ions as a result of prior interactions with rocks or minerals.

When a solution is in equilibrium with CO_2 gas in the stability field of magnetite, siderite may form from magnetite by an electrode half-reaction:



The two Fe^{3+} ions in magnetite must each accept one electron to convert them to the divalent state. This puts the electrons in equation 14.104 on the wrong side. Therefore, we reverse the equation and add CO_2 , H_2O , and H^+ as required for balance:



$$\Delta G_R^\circ = +12.225 \text{ kcal} \quad E^\circ = +0.265 \text{ V}$$

$$E_h = 0.265 + \frac{0.05916}{2} \log([\text{H}^+]^2[\text{CO}_2]^3) \quad (14.105)$$

$$E_h = 0.265 - 0.05916 \text{ pH} + 0.0887 \log [\text{CO}_2] \quad (14.106)$$

If $[\text{CO}_2] = 3 \times 10^{-4}$ atm (composition of the atmosphere):

$$E_h = -0.048 - 0.0592 \text{ pH} \quad (14.107)$$

Evidently, the siderite-magnetite boundary plots slightly *below* the stability limit of water when the solution is in contact with the atmosphere. Therefore, siderite is *not* expected to precipitate unless the fugacity of CO_2 is greater than 3×10^{-4} atm. If $[\text{CO}_2] = 10^{-2}$ atm, equation 14.106 yields:

$$E_h = +0.0876 - 0.0592 \text{ pH} \quad (14.108)$$

Under these conditions, siderite does have a small stability field and may precipitate under strongly reducing conditions from basic solutions as indicated in Figure 14.8

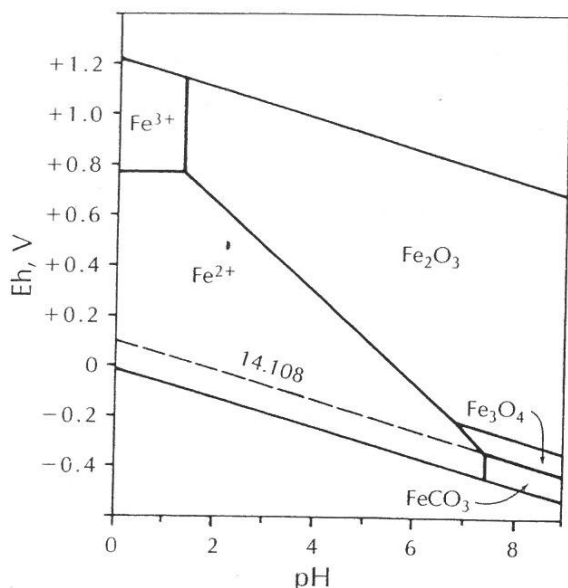


Figure 14.8 Stability of siderite, magnetite, and hematite in contact with water at 25°C in coordinates of Eh and pH assuming a fugacity of $\text{CO}_2 = 10^{-2}$ atm and activities of 10^{-6} mol/L for Fe^{2+} and Fe^{3+} .

However, the stability of siderite is restricted by its solubility in water under reducing conditions:



Since no oxidation or reduction takes place, equation 14.109 is based only on mass-balance considerations. At equilibrium:

$$\frac{[\text{CO}_2][\text{Fe}^{2+}]}{[\text{H}^+]^2} = K \quad (14.110)$$

Since $\Delta G_R^\circ = -9.491$ kcal and $K = 10^{6.958}$,

$$\log [\text{CO}_2] + \log [\text{Fe}^{2+}] - 2 \log [\text{H}^+] = 6.958 \quad (14.111)$$

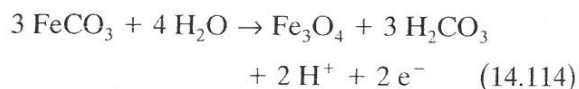
and

$$2 \text{pH} = 6.958 - \log [\text{CO}_2] - \log [\text{Fe}^{2+}] \quad (14.112)$$

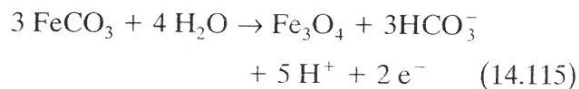
$$\text{If } [\text{CO}_2] = 10^{-2} \text{ mol/L and } [\text{Fe}^{2+}] = 10^{-6} \text{ mol/L,} \\ \text{pH} = 7.48 \quad (14.113)$$

Equation 14.113 limits the stability field of siderite in Figure 14.8, which also depicts the stability limits of hematite and magnetite in coordinates of Eh and pH at 25°C and a fugacity of $\text{CO}_2 = 10^{-2}$ atm.

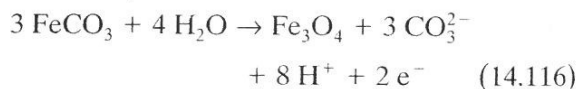
If a solution contains carbonate ions, but is not actually in contact with CO_2 gas, we must take into consideration the way in which the abundances of the carbonate species vary with the pH of the solution. Molecular carbonic acid dominates at $\text{pH} < 6.35$, bicarbonate dominates at $\text{pH} = 6.35$ – 10.3 , and the carbonate ion dominates at $\text{pH} > 10.3$ (Section 9.5, Table 9.3). Therefore, if the total carbonate content of a solution is 10^{-2} mol/L, $[\text{H}_2\text{CO}_3] = 10^{-2}$ mol/L at $\text{pH} < 6.35$, $[\text{HCO}_3^-] = 10^{-2}$ at $\text{pH} 6.35$ – 10.3 , and $[\text{CO}_3^{2-}] = 10^{-2}$ mol/L at $\text{pH} > 10.3$. The formation of siderite from magnetite in the different pH intervals can be expressed by electrode half-reactions modeled after equation 14.104. At $\text{pH} < 6.35$:



At $\text{pH} 6.35$ – 10.3 :



At $\text{pH} > 10.3$:



Similarly, the electrode half-reactions involving the conversion of siderite to hematite in the specified pH intervals are:

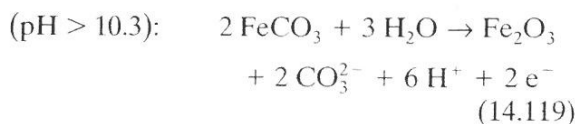
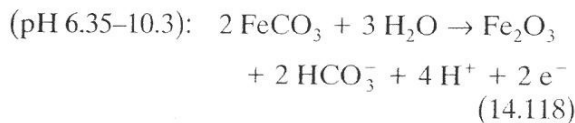
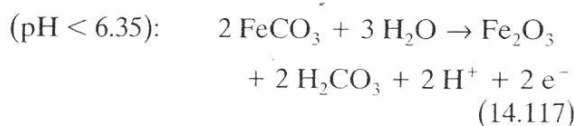


Table 14.4 Eh–pH Equations for the Conversion of Siderite to Magnetite and Hematite at Specified pH Intervals in Solutions Having a Total Carbonate Content of 10^{-2} mol/L Without Being in Contact with CO_2 Gas

A. Siderite to Magnetite

- pH < 6.35 (equation 14.114)
 $\text{Eh} = 0.217 - 0.0592 \text{ pH}$ (metastable) (1)
 pH 6.35–10.3 (equation 14.115)
 $\text{Eh} = 0.782 - 0.148 \text{ pH}$ (2)
 pH > 10.3 (equation 14.116)
 $\text{Eh} = 1.70 - 0.236 \text{ pH}$ (metastable) (3)

B. Siderite to Hematite

- pH < 6.35 (equation 14.117)
 $\text{Eh} = 0.211 - 0.0592 \text{ pH}$ (4)
 pH 6.35–10.3 (equation 14.118)
 $\text{Eh} = 0.587 - 0.118 \text{ pH}$ (5)
 pH > 10.3 (equation 14.119)
 $\text{Eh} = 1.20 - 0.177 \text{ pH}$ (metastable) (6)

The corresponding Eh–pH equations are listed in Table 14.4.

A comparison of equations 1 and 4 in Table 14.4 indicates that they have identical slopes but different intercepts, namely, 0.217 (siderite $\rightarrow$ magnetite) and 0.211 (siderite $\rightarrow$ hematite). This means that the conversion of siderite to magnetite (equation 1) is metastable because it occurs within the stability field of hematite where magnetite is not stable. For this reason, equation 4 in Table 14.4 forms the boundary between siderite and hematite at pH < 6.35 in Figure 14.9.

Above pH = 6.35, the transformation of siderite to hematite is represented by equation 5 in Table 14.4 until the boundary enters the magnetite field where equation 2 is valid. The siderite field delineated in this way actually pinches out at pH < 10.3 when equation 2 intersects the lower stability limit of water. Therefore, the reactions represented by equations 3 and 6 (valid at pH > 10.3) are also metastable and are not shown in Figure 14.9.

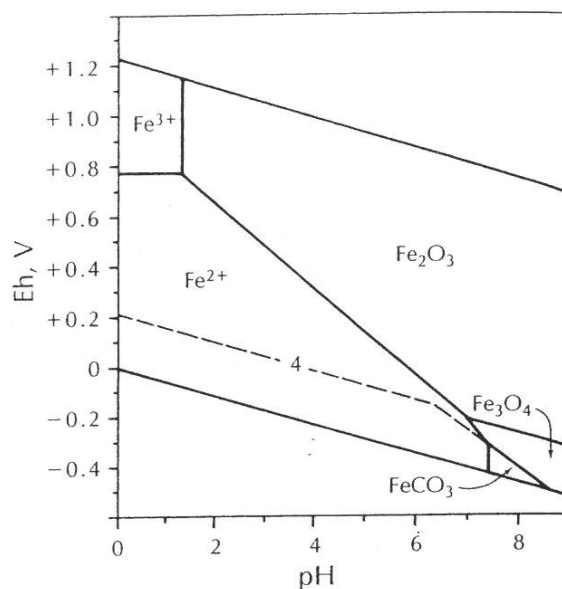


Figure 14.9 Stability of siderite in a solution containing 10^{-2} mol/L of dissolved carbonate ion at 25°C without being in contact with CO_2 gas. The dashed lines are extensions of equations 2, 4 and 5 in Table 14.4, which represent metastable transformations because at pH < 7.48 the solubility of siderite is greater than 10^{-6} mol/L.

The stability field of siderite delineated by these equations is restricted by the solubility of siderite with respect to Fe^{2+} expressed by equation 14.109. When the limit is taken to be $[\text{Fe}^{2+}] = 10^{-6}$ mol/L, the boundary between siderite and Fe^{2+} lies at pH = 7.48 (equation 14.113). The stability field of siderite *expands* as the total carbonate content of the system increases, but it *decreases* with increasing pH because the activity of CO_3^{2-} rises up to a pH value of about 10.5 and remains constant at higher pH values. When the system is in contact with CO_2 , the stability field of siderite increases with increasing fugacity of CO_2 and does not shrink as the pH rises (Figure 14.8) because, in this case, the activity of CO_3^{2-} increases with increasing pH. (See Problem 3 at the end of this chapter.) However, regardless of which case we consider (fixed carbonate content or CO_2 gas), siderite precipitates only when the carbonate content or the fugacity

of CO_2 are greater than is consistent with the present atmosphere of the Earth.

d. Stability of Sulfur Species

The dominant S-bearing ions and molecules in natural solutions are SO_4^{2-} , HSO_4^- , S^{2-} , HS^- , and H_2S^0 . The abundances of these ions in solutions containing a fixed amount of dissolved S depend both on pH and Eh because the transformation of sulfide to sulfate species requires a change in valence number from -2 to +6. In addition, native S (valence number = 0) may precipitate when the amount of S in the system exceeds its solubility.

The boundary between SO_4^{2-} and HSO_4^- can be calculated from the equilibrium:



$$\Delta G_R^\circ = -2.94 \text{ kcal} \quad K = 10^{2.155}$$

Therefore, when $[\text{HSO}_4^-] = [\text{SO}_4^{2-}]$, $[\text{H}^+] = 10^{-2.155} \text{ mol/L}$ and $\text{pH} = 2.16$. The boundaries between H_2S^0 , HS^- , and S^{2-} are also independent of Eh and occur at $\text{pH} = 7.0$ ($[\text{H}_2\text{S}] = [\text{HS}^-]$) and $\text{pH} = 12.9$ ($[\text{HS}^-] = [\text{S}^{2-}]$) as shown in Figure 9.3 (Section 9.5). Therefore, H_2S can be oxidized to HSO_4^- at $\text{pH} < 2.16$ and to SO_4^{2-} at $\text{pH} > 2.16$ but < 7.0 . In basic environments ($\text{pH} > 7.0$ but < 12.9) HS^- is oxidized to SO_4^{2-} , but S^{2-} becomes dominant only in extremely basic environments, which are rare in nature.

The electrode half-reactions that establish the boundaries between H_2S - HSO_4^- , H_2S - SO_4^{2-} , and HS^- - SO_4^{2-} are listed in Table 14.5A and have been plotted in Figure 14.10. The boundaries are drawn where the activities of reacting S species are equal to each other and are therefore independent of the total S content of the system.

Native S can precipitate from aqueous solution when the total S content of the system exceeds its solubility. This process is Eh-dependent because the valence numbers of S change from -2 to 0 and from +6 to 0 during the transformation of sulfide and sulfate to native S, respectively. Therefore, we establish the stability field of native S by a series of electrode half-reactions written for the Eh-pH conditions in which H_2S^0 , HS^- , HSO_4^- , and SO_4^{2-}

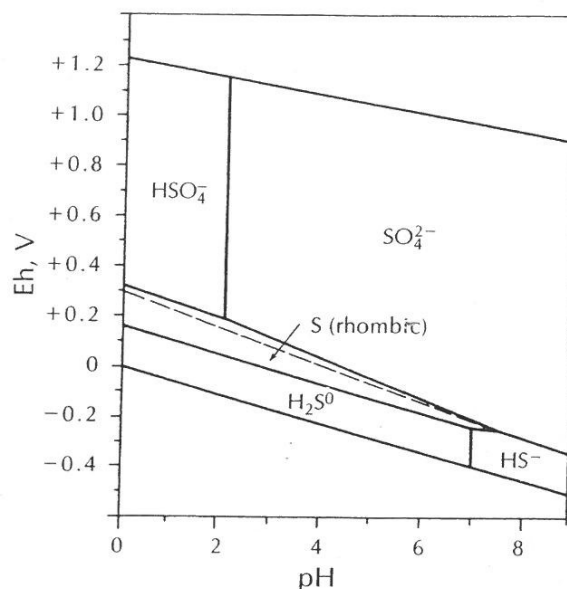


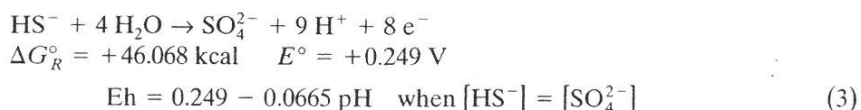
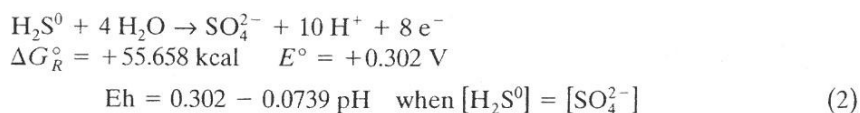
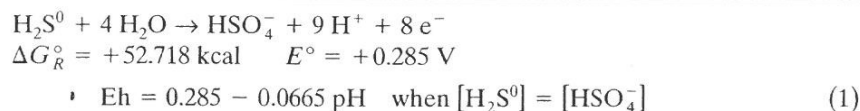
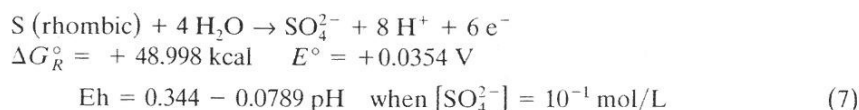
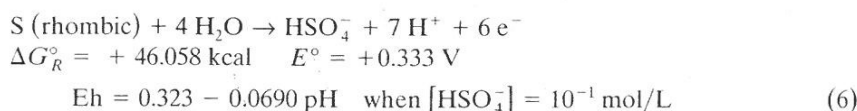
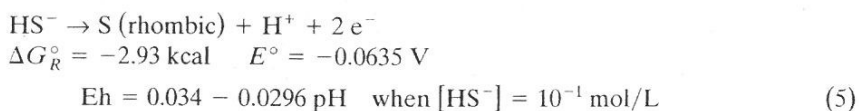
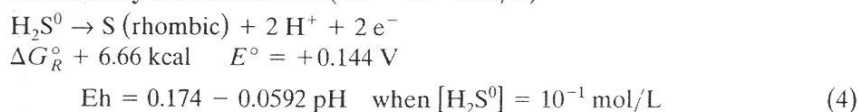
Figure 14.10 Abundance of aqueous S-species and stability field of native S (rhombic) in a system containing 10^{-1} mol/L of total S. The stability field of native S is based on the assumption that the aqueous species have activities of 10^{-1} mol/L in the Eh-pH fields in which they are the dominant ion or molecule.

dominate. Moreover, we specify that the total S content of the system is 10^{-1} mol/L in order to assure that the stability field of native S occupies a significant area in the diagram. The relevant equations are listed in Table 14.5B and have been used to draw the stability field of native S in Figure 14.10.

The diagram indicates that SO_4^{2-} and HSO_4^- are the dominant stable S species in well-oxygenated surface water, whereas H_2S^0 and HS^- predominate under reducing conditions corresponding to low O_2 fugacities. Native S has a wedge-shaped stability field that expands with the increasing total-S content of the system but disappears when $\Sigma \text{S} < 10^{-4.8} \text{ mol/L}$.

e. Stability of Ferrous Sulfides

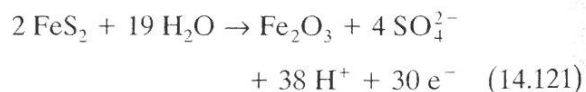
A comparison of Figures 14.7B and 14.10 indicates that the Eh-pH area in which Fe^{2+} is the dominant Fe ion in equilibrium with hematite and magnetite overlaps the fields of H_2S^0 ,

Table 14.5 Electrode Half-reactions Among Dissolved Molecular and Ionic S Species in Aqueous Solution at 25°C**A. Ionic and Molecular Species in Solution****B. Solubility of Native Sulfur ($\Sigma \text{S} = 10^{-1} \text{ mol/L}$)**

HSO_4^- , and SO_4^{2-} . Therefore, ferrous sulfides may be stable under a wide range of Eh and pH conditions, whereas ferric sulfides cannot form because Fe^{3+} is dominant *only* under highly oxidizing and acidic conditions in the area of HSO_4^- dominance.

The ferrous sulfides that may form include FeS (amorphous), pyrrhotite, mackinawite, troilite, pyrite, and marcasite. However, only pyrrhotite ($\text{Fe}_{0.95}\text{S}$) and pyrite (FeS_2) are stable and need to be considered here. The others may form as intermediate metastable products that recrystallize to form either pyrrhotite or pyrite in the course of geologic time.

In order to introduce ferrous sulfides into the Eh-pH diagram for the oxides of Fe, we guess that pyrite can occur with hematite in the SO_4^{2-} field. The transformation of pyrite into hematite plus SO_4^{2-} requires changes in the valence numbers of both Fe and S and therefore can be expressed as an electrode half-reaction:



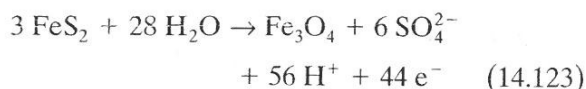
In this electrode, the valence numbers of S change from -1 in pyrite to $+6$ in sulfate, which causes the release of 28 electrons. In addition,

each of the Fe^{2+} ions in pyrite gives up one electron for a total of 30 electrons. For this reaction $\Delta G_R^\circ = +266.453 \text{ kcal}$ and $E^\circ = +0.385 \text{ V}$. If we set $[\text{SO}_4^{2-}] = 10^{-1} \text{ mol/L}$, the Eh-pH boundary equation is:

$$\text{Eh} = +0.377 - 0.0749 \text{ pH} \quad (14.122)$$

This line establishes the boundary between pyrite and hematite in that part of the stability field of hematite where SO_4^{2-} is dominant.

The further development of this Eh-pH diagram is best done by plotting each new line on Figure 14.11, which already contains the stability fields of the Fe-oxides as well as the distribution of S species. The pyrite-hematite boundary intersects the magnetite-hematite boundary, which implies that pyrite and magnetite can also coexist. Therefore, we construct a pyrite-magnetite electrode reaction in which the S appears as SO_4^{2-} :

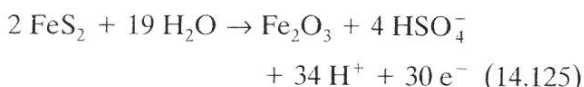


$$\Delta G_R^\circ = +395.136 \text{ kcal} \quad E^\circ = +0.389 \text{ V}$$

$$\text{Eh} = 0.381 - 0.0753 \text{ pH} \quad ([\text{SO}_4^{2-}] = 10^{-1.0} \text{ mol/L}) \quad (14.124)$$

By plotting equation 14.124 in Figure 14.11 we find that magnetite retains a small stability field in strongly reducing ($\text{Eh} < -0.30 \text{ V}$) and basic ($\text{pH} > 11.5$) environments where it can precipitate from sulfate-dominated solutions.

The pyrite-hematite boundary continues into the HSO_4^- field by virtue of the electrode half-reaction and the corresponding Eh-pH equation:



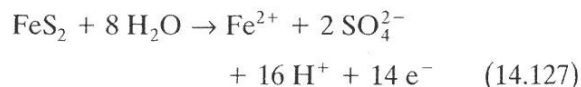
$$\Delta G_R^\circ + 254.693 \text{ kcal} \quad E^\circ = +0.368 \text{ V}$$

$$\text{Eh} = +0.360 - 0.0670 \text{ pH} \quad ([\text{HSO}_4^-] = 10^{-1} \text{ mol/L}) \quad (14.126)$$

Equation 14.126 extends the pyrite-hematite boundary from SO_4^{2-} -dominated environments into HSO_4^- -dominated environments with only a slight change in slope. However, we do not

plot equation 14.126 because we have entered the field where hematite is soluble with respect to Fe^{2+} .

Therefore, we must now consider the solubility of pyrite with respect to Fe^{2+} and the various S species. In the SO_4^{2-} -dominated Eh-pH region, the electrode-half reaction is:



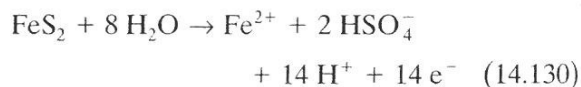
$$\Delta G_R^\circ = +118.146 \text{ kcal} \quad E^\circ = +0.366 \text{ V}$$

$$\text{Eh} = 0.366 + \frac{0.05916}{14} \times \log ([\text{H}^+]^{16} [\text{SO}_4^{2-}]^2 [\text{Fe}^{2+}]) \quad (14.128)$$

By converting $[\text{H}^+]$ to pH, setting $[\text{SO}_4^{2-}] = 10^{-1} \text{ mol/L}$ and $[\text{Fe}^{2+}] = 10^{-6} \text{ mol/L}$, equation 14.128 reduces to:

$$\text{Eh} = 0.322 - 0.067 \text{ pH} \quad (14.129)$$

In the HSO_4^- -dominated field the electrode half-reaction is:



$$\Delta G_R^\circ = 112.266 \text{ kcal} \quad E^\circ = +0.348 \text{ V}$$

$$\text{Eh} = 0.316 - 0.0592 \text{ pH} \quad (14.131)$$

where $[\text{HSO}_4^-] = 10^{-1} \text{ mol/L}$ and $[\text{Fe}^{2+}] = 10^{-6} \text{ mol/L}$.

Because this line (equation 14.131) enters the stability field of native S at a very low angle we must consider the solubility of pyrite with respect to Fe^{2+} and native S.



$$\Delta G_R^\circ = +20.15 \text{ kcal} \quad E^\circ = +0.437 \text{ V}$$

$$\text{Eh} = +0.437 + \frac{0.05916}{2} \log [\text{Fe}^{2+}] \quad (14.133)$$

If $[\text{Fe}^{2+}] = 10^{-6} \text{ mol/L}$, $\text{Eh} = 0.259 \text{ V}$.

The construction of Figure 14.11 indicates that pyrite is stable in the Eh-pH region lying below the boundaries based on equations 14.122, 14.124, 14.129, 14.131, and 14.133. Therefore,

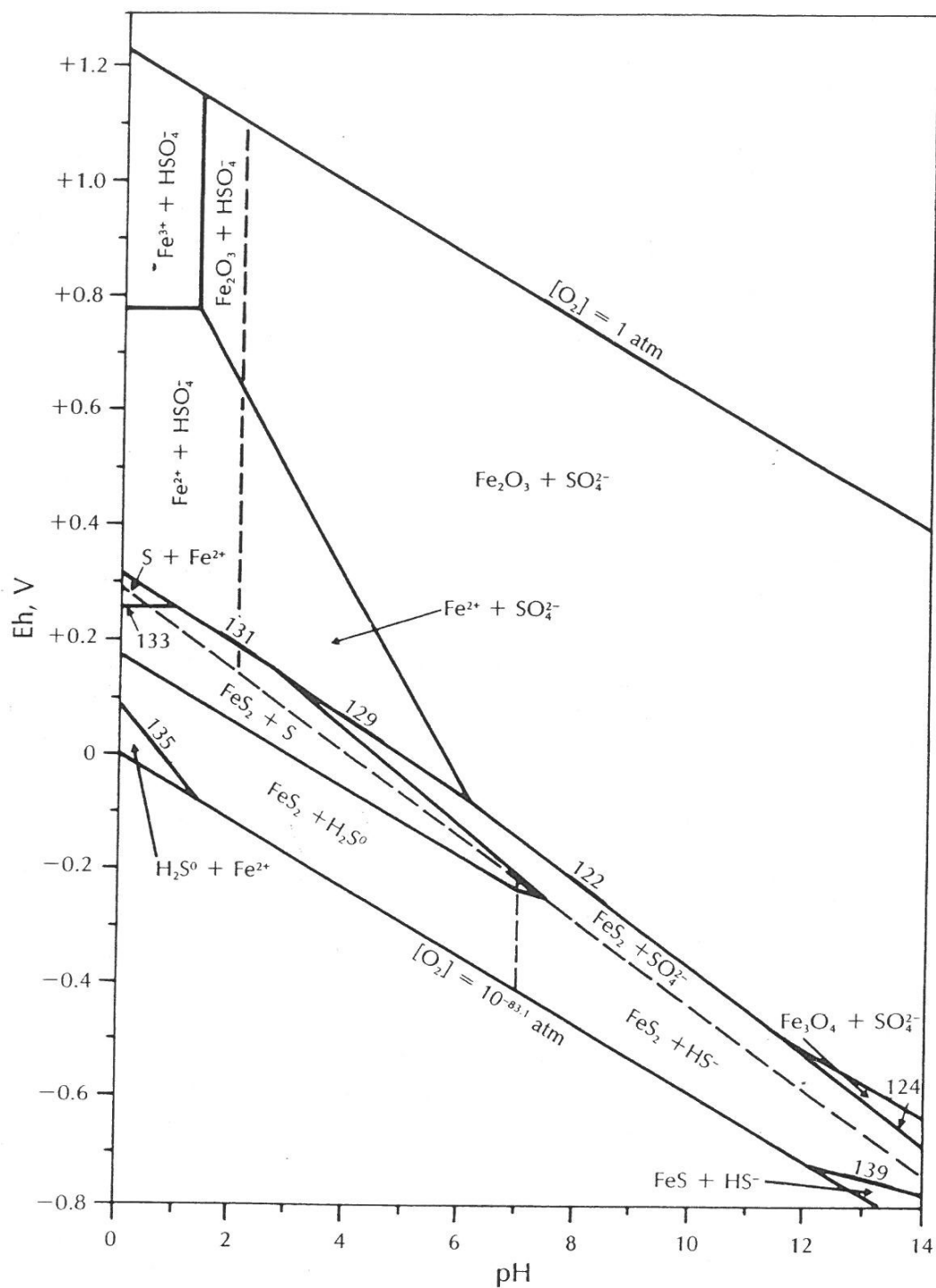
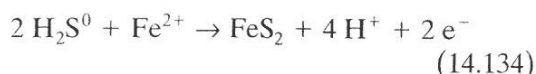


Figure 14.11 Composite Eh-pH diagram for Fe oxides and sulfides in the presence of water at 25°C containing $\Sigma S = 10^{-1}$ mol/L. In addition, the dashed lines indicate the conditions under which the different S species become dominant. The numbers adjacent to some of the boundaries identify equations in the text. An even more comprehensive diagram including oxides, sulfides, and the carbonate of Fe was published by Garrels and Christ (1965).

pyrite occurs in the area where H_2S^0 is dominant and could be soluble with respect to Fe^{2+} and H_2S^0 . The relevant electrode half-reaction is:

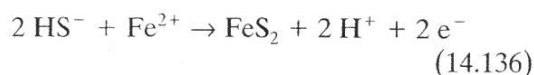


$$\Delta G_R^\circ = -6.83 \text{ kcal} \quad E^\circ = -0.148 \text{ V}$$

$$\text{Eh} = +0.088 - 0.118 \text{ pH} \quad (14.135)$$

where $[\text{H}_2\text{S}^0] = 10^{-1} \text{ mol/L}$ and $[\text{Fe}^{2+}] = 10^{-6} \text{ mol/L}$.

Since pyrite also occurs where HS^- is the dominant S species, we write the electrode half-reaction:

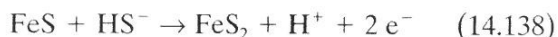


$$\Delta G_R^\circ = 26.01 \text{ kcal} \quad E^\circ = -0.564 \text{ V}$$

$$\text{Eh} = -0.327 - 0.0591 \text{ pH} \quad (14.137)$$

where $[\text{HS}^-] = 10^{-1} \text{ mol/L}$ and $[\text{Fe}^{2+}] = 10^{-6} \text{ mol/L}$. Evidently this line lies outside the stability field of water, which implies that pyrite is not appreciably soluble with respect to Fe^{2+} under conditions where HS^- is the dominant S species.

Finally, we consider the relationship between pyrite and pyrrhotite, which does require a change in the valence number of S from -1 in pyrite to -2 in pyrrhotite. The relevant electrode reaction in the HS^- -dominated Eh-pH region is:



$$\Delta G_R^\circ = -17.93 \text{ kcal} \quad E^\circ = -0.389 \text{ V}$$

$$\text{Eh} = -0.359 - 0.0296 \text{ pH}$$

$$([\text{HS}^-] = 10^{-1} \text{ mol/L}) \quad (14.139)$$

Note that the formula for pyrrhotite in equation 14.138 is given as FeS for the sake of simplicity even though its standard free energy value in Appendix B is for $\text{Fe}_{0.95}\text{S}$. Equation 14.139 delineates a small stability field for pyrrhotite under extremely basic ($\text{pH} > 12$) and reducing ($\text{Eh} < -0.52 \text{ V}$) conditions, which explains why this mineral is rarely, if ever, found as an authigenic mineral in sedimentary rocks.

Figure 14.11 contains a great deal of information about the geochemistry of Fe on the surface of the Earth. It identifies the environments in which hematite, magnetite, pyrite, pyrrhotite, and native S are stable in the presence of water containing 10^{-1} mol/L of S species. In addition, the diagram delineates environmental conditions that cause the Fe-minerals to dissolve significantly to form Fe^{2+} or Fe^{3+} and which therefore favor the transport of Fe in aqueous solutions at 25°C . In general, we see that Fe is mobile primarily under acidic and oxidizing conditions and that it becomes immobile when Fe-bearing solutions are neutralized or enter reducing environments containing dissolved S. The minerals that form by precipitation from aqueous solutions depend on the Eh and pH conditions as shown in Figure 14.11. However, intermediate compounds may form initially, which may subsequently recrystallize to form the stable phases having the lowest standard free energies of formation.

All natural solutions contain not only Fe and S but also carbonate ions. Therefore, natural systems may precipitate siderite, which is not shown in Figure 14.11 in order to avoid overcrowding. An Eh-pH diagram including the oxides, sulfides, and the carbonate of Fe was worked out by Garrels and Christ (1965, Fig. 7.21) for a system at 25°C containing $\Sigma\text{S} = 10^{-6} \text{ mol/L}$ and $\Sigma\text{carbonates} = 1 \text{ mol/L}$.

The methods illustrated here can be used to develop Eh-pH diagrams for all elements in the periodic table with the exception of the noble gases, which do not change their oxidation state. Such diagrams are useful in the study of metallic mineral deposits, in geochemical prospecting, and in environmental geochemistry. The most comprehensive treatment of this subject is by Brookins (1988).

14.6 Summary

The elements and their ions can be ranked by their tendency to give up electrons as expressed in displacement reactions that involve the transfer of electrons from the electron donor (reducing

agent) to the electron acceptor (oxidizing agent). The resulting oxidation-reduction reactions must be balanced by use of valence numbers to assure that the number of electrons given up by the reducing agent is equal to the number of electrons gained by the oxidizing agent.

Oxidation-reduction reactions can be divided into two complementary half-reactions that may be separated physically to form an electrochemical cell. The electromotive force (emf) exerted by such a cell is measurable as a voltage between the electrodes dipping into an electrolyte solution. The emf of an electrochemical cell in the standard state is related to the standard free energy change of the reaction by Faraday's constant. When the cell is *not* in the standard state, the emf is related to the reaction quotient and decreases to zero as the oxidation-reduction reaction approaches equilibrium.

A systematic scale of emf values is created by arbitrarily assigning a value of zero to the emf of the H_2 electrode in the standard state. This is accomplished by setting the standard free energies of formation of H^+ and e^- equal to zero. The standard electrode potentials of all other electrodes can be calculated from their standard free energy change alone when they are connected to the standard H_2 electrode. When the electrode half-reactions are written with the electrons on the right side of the equation, the algebraic sign and magnitude of their standard electrode potentials become the basis for ranking them relative to the standard H_2 electrode. Such a ranked list of electrode half-reactions is called the electromotive series. A negative sign of the standard electrode potential indicates that the half-reaction is a stronger electron donor (reducing agent) than the standard H_2 electrode, whereas a positive sign indicates that the electrode is a weaker reducing agent and is therefore a stronger oxidizing agent than the standard H_2 electrode.

As expected from their electronic configurations, the alkali metals and alkaline earths are strong reducing agents, whereas the halogens and other nonmetallic elements are strong oxidizing agents. However, the cations of a metal can act as

electron acceptors (oxidizing agents) and the anions of a nonmetal can function as electron donors (reducing agents).

In order to facilitate the application of electrochemistry to natural environments a special emf is defined as the voltage generated by any electrode in any state relative to the H_2 electrode in the standard state (Eh). Natural environments on the surface of the Earth are restricted to a range of Eh and pH conditions determined by the dissociation of liquid water into H^+ and O_2 , which releases electrons and therefore constitutes an electrode half-reaction. The limits on the stability of water result from the requirement that the fugacity of O_2 in an environment open to the atmosphere must be less than 1.0 but greater than $10^{-83.1}$ atm. The Eh and pH values of natural environments depend on their chemical compositions and on the reactions that are taking place within them. Therefore, the Eh and the pH are both *environmental parameters* that are imposed on any specific reaction occurring in the environment.

The algebraic sign of measured environmental Eh values gives meaning to the widely used concept of *oxidizing* and *reducing* conditions. Oxidizing environments have positive Eh values, so they are weaker reducing agents than the standard H_2 electrode. Reducing environments have negative Eh values and are therefore stronger reducing agents than the standard H_2 electrode. The capacity of an environment to be a reducing agent results from the presence of potential electron donors that can force the standard H_2 electrode to accept electrons. Organic compounds of biological origin containing reduced forms of C, such as CH_4 , where the C has a valence number of -4 , are a common source of electrons in natural environments. Similarly, the oxidizing character of an environment results from the presence of potential electron acceptors or oxidizing agents. A common oxidizing agent in natural environments is molecular O_2 , which can accept four electrons per molecule. However, many other compounds and ions have an effect on the Eh of a particular environment.

The Eh and pH of natural environments can be used to define relationships between solids,

between solids and ions, and among ions in aqueous solution of elements that can have different valence numbers. The Eh–pH stability diagrams of Fe compounds, including oxides, carbonates, and sulfides, are constructed from appropriate electrode half-reactions that define compatible relationships between Eh and pH for these electrodes. The resulting diagrams can be used to understand the geochemistry of Fe and S and to predict the

effect of environmental changes on the compounds and ions of these elements. Such diagrams have been constructed for most of the elements. The Eh–pH diagrams are useful in studying the origin of metallic mineral deposits and the subsequent dispersal of metals during chemical weathering. For this reason they are relevant also in hydrogeology, environmental geochemistry, and soil chemistry.

Problems

- Write electronic formulas for S^{2-} , Ti^{4+} , P^{3-} , Zr^{4+} , and Cl^{7+} . Note the tendency of these elements to achieve the electronic configuration of noble gases.
- Balance the following oxidation–reduction reactions.
 - $MnO_4^- + Cl^- \rightarrow Mn^{2+} + Cl_2$
 - $As_2S_3 + NO_3^- \rightarrow HAsO_3 + S + NO$
 - $Cr_2O_7^{2-} + I^- \rightarrow Cr^{3+} + I_2$
 - $CrO_2^- + ClO^- \rightarrow CrO_4^{2-} + Cl^-$
 - $CH_4^0 + SO_4^{2-} \rightarrow HCO_3^- + HS^-$
- Calculate the abundance of CO_3^{2-} in moles per liter at pH = 9, 10, 11 in two different systems.
 - Having a fixed total carbonate content of 10^{-2} mol/L.
 - In equilibrium with CO_2 gas at a fugacity of 10^{-2} atm.
- Complete the following electrode half-reactions in the conventional form and calculate their standard electrode potentials.
 - $PbO \text{ (red)} \rightarrow PbO_2$
 - $Cu_2O \rightarrow Cu$
 - $UO_2^{2+} \rightarrow U^{4+}$
 - $Cr^{3+} \rightarrow Cr_2O_7^{2-}$
 - $Si \rightarrow SiF_6^{2-}$
- Combine the Ag and Cu electrode half-reactions and calculate the emf when $[Cu^{2+}]/[Ag^+]^2 = 10^{-4}$.
- Combine the Al and Ni electrodes and calculate the emf when $[Ni^{2+}]^3/[Al^{3+}]^2 = 10^{-2}$.
- Manganese forms the following oxide minerals: manganite ($MnOOH$), manganosite (MnO), pyrolusite (MnO_2), and hausmannite (Mn_3O_4). Determine the valence numbers of Mn in each compound and order them in terms of increasing valence number of Mn.
 - Derive Eh–pH equations based on the progressive oxidation of Mn to form oxides starting with metallic Mn.
 - Plot these equations in coordinates of Eh and pH using a full sheet of graph paper, add the stability limits of water, and label all fields.
 - Is metallic Mn stable at the surface of the Earth?
- Derive equations in terms of Eh and pH for the solution of the oxide minerals of Mn to form Mn^{2+} .
 - Plot the equations for $[Mn^{2+}] = 10^{-6}$ mol/L on the diagram constructed in Problem 8b.
- What is the fugacity of O_2 of an environment having pH = 7.0 and Eh = +0.2 V?
- A solution containing 10^{-6} mol/L of Fe^{2+} and Mn^{2+} at Eh = +0.2 V is neutralized such that its pH changes from 4 to 6. What effect will this have on the concentrations of Fe^{2+} and Mn^{2+} in the solution? Refer to the appropriate Eh–pH diagrams.
- Based on the insight gained above, explain why Fe-oxide deposits in near shore marine environments have low Mn concentrations and speculate about the fate of Mn in the oceans.

References

- BAAS-BECKING, L. G. M., I. R. KAPLAN, and O. MOORE, 1960. Limits of the natural environment in terms of pH and oxidation-reduction potentials. *J. Geol.*, **68**:243-284.
- BROOKINS, D. G., 1978. Eh-pH diagrams for elements from $Z = 40$ to $Z = 52$: Application to the Oklo natural reactor. *Chem. Geol.*, **23**:324-342.
- BROOKINS, D. G., 1988. *Eh-pH Diagrams in Geochemistry*. Springer-Verlag, New York, 200 pp.
- GARRELS, R. M., 1960. *Mineral Equilibria*. Harper & Row, New York, 244 pp.
- GARRELS, R. M., and C. L. CHRIST, 1965. *Solutions, Minerals and Equilibria*. Harper & Row, New York, 450 pp.
- HOGNESS, T. R., and W. C. JOHNSON, 1954. *Qualitative Analysis and Chemical Equilibrium*, 4th ed. Henry Holt, New York, 621 pp.
- POURBAIX, M. J. N., 1949. *Thermodynamics of Dilute Aqueous Solutions*. Arnold, London.
- STUMM, W., and J. J. MORGAN, 1996. *Aquatic Chemistry*, 3rd ed. Wiley, New York, 1022 pp.

Rates of Geochemical Processes

The progress of chemical reactions depends on their *rates*, which are controlled by a variety of factors, including the temperature, the concentrations of reactants and products, and the rate of mass transfer into and out of a fixed reaction site. In some cases, the transformation that we observe is the end product of several intermediate reactions each of which proceeds at its own rate. As a result, the rate of the complete reaction may be controlled by the slowest of the intermediate reactions, which becomes the rate-determining step. Therefore, the study of the kinetics of chemical reactions leads to insights about what *really* happens during such familiar processes as the formation of a crystalline precipitate in aqueous solution, the incongruent solution of minerals, or the formation of concretions during diagenesis of sediment. The explanations of *reaction mechanisms* and the mathematical descriptions of the rates of chemical transformations are often complex and are, in many cases, not as well understood as the thermodynamic treatment of chemical reactions. In general, thermodynamics allows us to predict what *should* happen, whereas kinetics determines whether it actually does occur in a reasonable interval of time.

15.1 Rates of Chemical Reactions

The first quantitative study of chemical kinetics was undertaken by L. Wilhelmy in 1850. It was followed in 1863 by the important work of

Guldberg and Waage, who derived the Law of Mass Action from the concept that at equilibrium the rate of the forward reaction is equal to the rate of the backward reaction. The principles of chemical kinetics and other rate-dependent processes such as diffusion and advection have been combined with thermodynamics only recently in comprehensive mathematical descriptions of the rates of geochemical processes (Berner, 1971, 1980; Hofmann et al., 1974; Lerman, 1979; Lasaga and Kirkpatrick, 1981).

The rate of a chemical reaction can be expressed in terms of the consumption of one of the reactants or in terms of the formation of one of the products. Whatever method is used, the rates of chemical reactions *decline* with time from some initial value until they become virtually equal to zero as the reaction is either reaching *equilibrium* or is going to *completion*. We first encountered this idea in Section 9.1, where we described the dissolution of NaCl(s) in water as a function of time (Figure 9.1).

a. Order of the Reaction

The rate of change of the concentration of a reactant or product of a chemical reaction generally varies as some power of its concentration. For example, in the reaction:



the rate of consumption of reactant X is given by an equation of the form:

$$-\frac{d(X)}{dt} = k(X)^n \quad (15.2)$$

where (X) is the concentration of reactant X at any time t , n is a number whose value indicates the *order* of the reaction, k is the specific rate constant, and the minus sign is needed to indicate that the rate *decreases* with time. The numerical values of n and k for a specific reaction must be determined experimentally. If $n = 1.0$, the reaction is said to be a *first-order* reaction, and so on. The value of n may also be a fraction like $\frac{1}{2}$ in which case the order is *one half*.

The *order* of the reaction represented by equation 15.1 can be determined by measuring how the *initial* rate of the reaction (R_0^X) varies with the initial concentration of reactant X represented by $(X)_0$. For this purpose we rewrite equation 15.2 as:

$$R_0^X = \text{constant} \times ((X)_0)^n \quad (15.3)$$

which we linearize in the form:

$$\log R_0^X = \log \text{constant} + n \log (X)_0 \quad (15.4)$$

Therefore, a plot of $\log R_0^X$ versus $\log (X)_0$ yields a straight line whose slope is equal to n .

The rate equation of reactions such as:



may take the form:

$$-\frac{d(X)}{dt} = k(X)^2 \quad (15.6)$$

which makes it a second-order reaction. If the reaction is:



the rate equation may take the form:

$$-\frac{d(X)}{dt} = k(X)^1(Y)^1 \quad (15.8)$$

This is also a second-order reaction, but it is said to be first order with respect to X , first order with respect to Y , and second order overall. In general, the rate equation of a chemical reaction:



may be expressed as:

$$-\frac{d(X)}{dt} = k(X)^{n_1}(Y)^{n_2} \quad (15.10)$$

and the order of such a reaction is $n_1 + n_2$. Note that the exponents of the rate equation *do not* necessarily agree with the molar coefficients of the balanced equation of the reaction.

b. Rate Constant

The value of the specific rate constant (k) must also be determined experimentally by measuring the concentration of reactant X at known intervals of time as the reaction runs its course. For this purpose we rearrange equation 15.2 for a first-order reaction:

$$-\frac{d(X)}{(X)} = kdt \quad (15.11)$$

which we can integrate:

$$\int \frac{d(X)}{(X)} = -k \int dt \quad (15.12)$$

$$\ln(X) = -kt + \text{constant of integration} \quad (15.13)$$

If $t = 0$, the integration constant is $\ln(X)_0$ where $(X)_0$ is the initial concentration of X . Therefore,

$$\ln(X) = \ln(X)_0 - kt \quad (15.14)$$

$$\ln \frac{(X)}{(X)_0} = -kt \quad (15.15)$$

$$(X) = (X)_0 e^{-kt} \quad (15.16)$$

Equation 15.14 is a straight line in coordinates of $\ln(X)$ and t whose slope is negative because the concentration of reactant X *decreases* with time. The absolute value of the slope is the specific rate constant (k) of the reaction. Equation 15.16 has the same form as equation 16.9 in Chapter 16, which expresses the *law of radioactive decay*.

The rate of decay of a *radioactive* nuclide can be stated in terms of the *half-life*, which is the time required for one half of an initial number of radioactive atoms to decay. We therefore define the *reaction half-life* ($t_{1/2}$) in a similar manner and

can show (see Chapter 16) that for a first-order reaction:

$$t_{1/2} = \frac{\ln 2}{k} \quad (15.17)$$

When time is measured in seconds and concentrations are expressed in moles/liter, the units of the rate constant (k) of a first-order reaction are:

$$k = \frac{d(X)/dt}{(X)} = \frac{\text{mol/L}}{\text{mol/L} \times \text{sec}} = \text{sec}^{-1} \quad (15.18)$$

For second-order reactions the units of k are:

$$\begin{aligned} k &= \frac{d(X)dt}{(X)^2} = \frac{\text{mol/L}}{(\text{mol/L})^2 \times \text{sec}} \\ &= \frac{\text{L}}{\text{mol} \times \text{sec}} = \text{L mol}^{-1} \text{sec}^{-1} \quad (15.19) \end{aligned}$$

The rate of a first-order reaction may also be expressed in terms of the *increase* of the concentration of one of the *products*. For example, for the reaction represented by equation 15.1 the rate of formation of product Y can be expressed as follows. If x moles of reactant X have reacted at time t , the concentration of products Y and Z is equal to x and the concentration of reactant X is $(X)_0 - x$. Therefore, the rate of formation of Y is:

$$\frac{dx}{dt} = k[(X)_0 - x] \quad (15.20)$$

By rearranging this equation and by integration we obtain:

$$-\ln [(X)_0 - x] = kt + \text{constant} \quad (15.21)$$

If $t = 0$, then $x = 0$, and the constant of integration is $-\ln (X)_0$. Therefore:

$$\ln \left(\frac{(X)_0}{(X)_0 - x} \right) = kt \quad (15.22)$$

$$\frac{(X)_0}{(X)_0 - x} = e^{kt} \quad (15.23)$$

$$\frac{(X)_0 - x}{(X)_0} = e^{-kt} \quad (15.24)$$

and thus

$$x = (X)_0(1 - e^{-kt}) \quad (15.25)$$

Equation 15.22 is a straight line in coordinates of $\ln (X)_0 / [(X)_0 - x]$ and t whose slope is equal to k .

Equation 15.25 has the same form as equation 16.19 for the growth of a stable radiogenic daughter. Therefore, the consumption of a reactant and the formation of a product in a *first-order* chemical reaction are governed by the same equations as the decay of a radioactive parent nuclide and the resulting growth of a radiogenic daughter. Both are illustrated in Figure 15.1 for a hypothetical reaction whose rate constant $k = 0.0693 \text{ sec}^{-1}$, which corresponds to a reaction half-life of 10 sec.

c. Temperature Dependence of Rate Constants

Before a reaction of the form:



can take place, the atoms or molecules of the reactants X and Y must collide with *sufficient* energy to allow them to form a bond. The minimum energy required for two atoms or molecules to react is called the *activation energy*. When X and Y are molecules, the chemical bonds they contain may have to break before the atoms can recombine to form an intermediate *activated complex*, which subsequently decomposes to form the product. The activation energy (E_a) acts as a barrier, illustrated in Figure 15.2, between the reactants and products that must be overcome before the reactants can form the desired product. The rates of chemical reaction ultimately depend on the *frequency* with which atoms or molecules of the reactants acquire the necessary activation energy that enables them to form bonds. Therefore, the rates of chemical reactions *increase* with increasing temperature because the fraction of atomic or

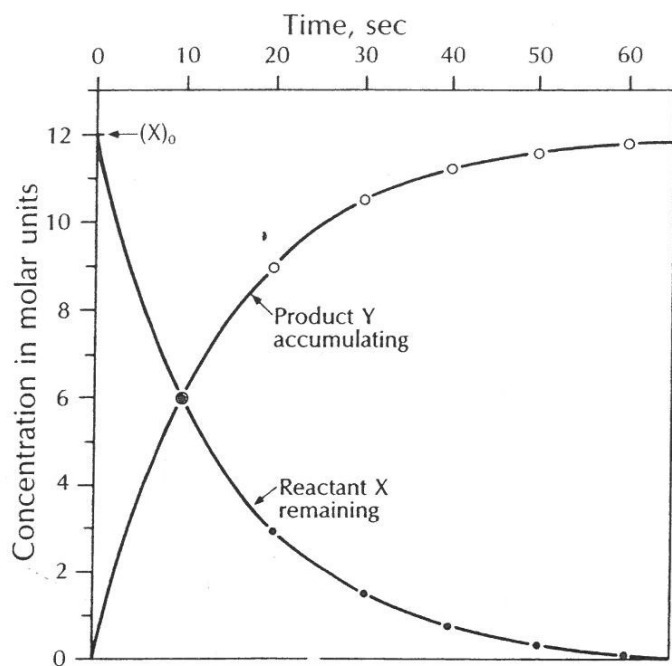


Figure 15.1 Kinetics of a hypothetical first-order chemical reaction represented by $X \rightarrow Y + Z$. The consumption of reactant X follows equation 15.16 where the specific rate constant $k = 0.0693 \text{ sec}^{-1}$, corresponding to a reaction half-life of 10 sec. The formation of products Y and Z is given by equation 15.25. The graph shows that the rate of consumption of reactant X and the rate of formation of product Y both *decrease* with time from their initial values at $t = 0$. In this example the concentration of reactant X approaches zero, indicating that the reaction is going to *completion*. If a finite amount of X remains when the reaction rates approach zero, then the reaction is in *equilibrium*. Note that both curves represent the forward reaction. The consumption of a reactant and the formation of a product in a first-order chemical reaction are governed by the same equations as the decay of a radioactive parent nuclide to a stable radiogenic daughter discussed in Chapter 16.

molecular collisions that are sufficiently energetic to provide the activation energy increases with temperature.

The relationship between the specific rate constant (k) of a chemical reaction, the temperature (T) in kelvins, and the activation energy (E_a) in kilocalories is given by an equation derived by Svante Arrhenius in 1889 (Moore, 1955):

$$k = Ae^{-E_a/RT} \quad (15.27)$$

where A is a constant that has a characteristic value for each reaction, and R is the gas constant. The constant A is called the *frequency factor* and $e^{-E_a/RT}$ is known as the *Boltzmann factor*, which is the fraction of atoms or molecules that manages to acquire the necessary activation energy from the energy liberated by collisions. By taking base 10 logarithms of equation 15.27 we can transform it into a straight line in coordinates of $\log k$ and $1/T$:

$$\log k = \log A - \frac{E_a}{2.303R} \left(\frac{1}{T} \right) \quad (15.28)$$

A plot of experimentally determined values of the rate constant of a chemical reaction at differ-

ent temperatures should therefore yield a straight line whose slope is equal to:

$$m = -\frac{E_a}{2.303R} \quad (15.29)$$

from which the activation energy of the reaction can be calculated. The intercept on the $\log k$ axis is equal to the logarithm of the frequency factor A . The activation energy is itself slightly dependent on temperature, but the slope of equation 15.28 is not perceptibly changed by this effect.

The formation of the activated complex and its deactivation to form the product of a chemical reaction each have their own rate constants. As a result, even a straightforward first-order reaction involves internal mechanisms that are difficult to express mathematically. The problem can be solved only for the condition of *steady state* when the rate of formation of the activated complex is equal to the rate of its deactivation and the number of activated complexes remains constant. Although the concept of the activated complex was initially developed for reactions in the gas phase, it also applies to reactions among ions in aqueous solution.

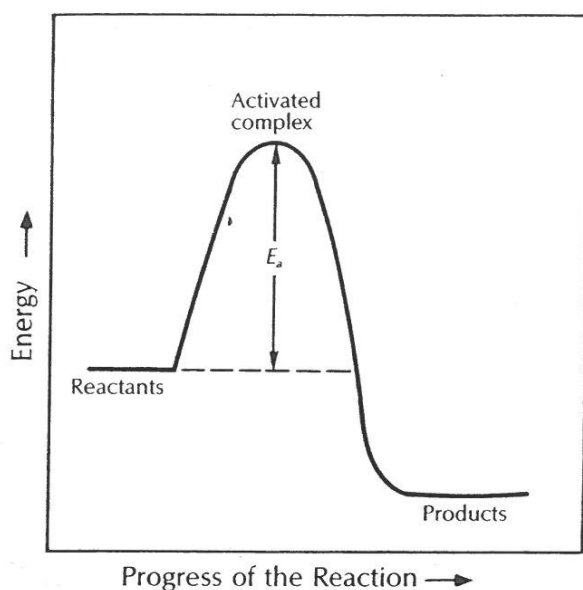


Figure 15.2 Schematic representation of the activation energy (E_a) that must be acquired by the reactants of a chemical reaction to form the activated complex, which subsequently decomposes to form the products.

15.2 Transport of Matter: Advection

The equations describing the rates of chemical reactions discussed above apply to stationary systems such as a cylinder of gas or an aqueous electrolyte solution in a beaker. However, chemical reactions in nature commonly require the movement of reactants or products and therefore depend on the applicable transport mechanism.

Transport of a solute in a solution can occur primarily by advection or by diffusion. *Advection* involves the displacement of matter in response to the action of a force. Examples of advection include the flow of water in the channel of a stream or in subsurface in response to a hydraulic gradient. In addition, rain falling from a cloud, sediment settling out of water, and the movement of lithospheric plates on the surface of the Earth are all examples of advection.

The displacement of matter produces a *flux* (F), which is the product of the density of the material (d) and its velocity (v):

$$F = dv \quad (15.30)$$

When the density is expressed in grams per cubic centimeter and the velocity in centimeters per second, the units of the flux are:

$$\begin{aligned} \text{units of flux } (F) &= \frac{\text{g} \times \text{cm}}{\text{cm}^3 \times \text{sec}} \\ &= \text{g cm}^{-2} \text{sec}^{-1} \quad (15.31) \end{aligned}$$

When the material in motion contains another substance X in solution at a concentration (X) expressed in grams per cubic centimeter, then the flux of X is:

$$F_X = (X)v \quad \text{g cm}^{-2} \text{sec}^{-1} \quad (15.32)$$

In general, typical velocities of advective transport processes are (Lerman, 1979):

	Meters per year
Particles settling in water	3×10^2 to 3×10^4
Groundwater flow	1–100
Upwelling in the oceans	1–2
Sedimentation in oceans and lakes	3×10^{-6} to 1×10^{-6}

The flow of water through the pore spaces of sediment or rocks in the subsurface (i.e., an aquifer) is an important advective transport mechanism that affects chemical reactions occurring within the aquifer because the water may supply some of the reactants or remove some of the products. The movement of water, or of another liquid such as petroleum, through a porous medium is governed by a law that was stated by Henry Darcy in 1856. *Darcy's law* expresses the flux of water (F) flowing through unit area (a) of an aquifer in unit time as a function of the hydrostatic pressure difference, which is equivalent to the hydraulic gradient (S):

$$F = KS \quad (15.33)$$

where K is the *coefficient of hydraulic conductivity* whose dimensions are centimeters per second. The flux F has the dimensions of volume/unit area/unit time ($\text{cm}^3 \text{cm}^{-2} \text{sec}^{-1}$), which reduces to

centimeter per second and is therefore equivalent to the velocity of the water. The hydraulic gradient S is given by:

$$S = \frac{\Delta h}{L} \quad (15.34)$$

where Δh is the difference in elevation between two points within the aquifer and L is the distance between these points measured along the flow path. Equation 15.33 can be modified to yield the *discharge* (D) in units of cubic centimeters per second:

$$D = K S a = F a \quad (15.35)$$

where a is the cross-sectional area of the aquifer shown schematically in Figure 15.3. When the dip of the aquifer is zero ($\Delta h = 0$), flow occurs only when a pressure difference is imposed on the aquifer.

The hydraulic conductivity (K) is related to the *coefficient of hydraulic permeability* (k) by the equation:

$$K = \frac{k d g}{\eta} \quad (15.36)$$

where g is the acceleration due to gravity and η is the viscosity of the fluid. For water at 20°C, $\eta = 0.01 \text{ g cm}^{-1} \text{ sec}^{-1}$, $d = 1.0 \text{ g/cm}^3$, and since $g = 980 \text{ cm/sec}^2$:

$$K \approx 10^5 k \quad (15.37)$$

The coefficient of hydraulic permeability has the dimension of length squared and is related to the

porosity of the aquifer and to the radius of the particles of which it is composed (Lerman, 1979). Values of k range over 12 orders of magnitude from 10^{-3} cm^2 for clean gravel to 10^{-15} cm^2 for granite and unweathered clay. The hydraulic conductivity (K) varies similarly. Good aquifers composed of clean sand or sand-gravel mixtures characteristically have $K \approx 10^{-1} \text{ cm/sec}$ and $k \approx 10^{-6} \text{ cm}^2$ (Lerman, 1979).

We can now consider an aquifer within which a chemical reaction is occurring (Moore, 1955, p. 544). A solution containing reactant X at concentration (X) mol/L flows through a volume element dV (the reactor) of the aquifer with a discharge rate D (L/sec) and leaves the reactor with a concentration (X) $- d(X)$. If no mixing occurs, then the net change in the number of moles of X with time in the reactor is the sum of two terms. One of these is the effect of the chemical reaction in the reactor, and the other is the difference in the concentration of X between the inflow and outflow. Hence:

$$\frac{dn_x}{dt} = R dV - D d(X) \quad (15.38)$$

where R is the reaction rate per unit volume. For a first-order reaction with respect to reactant X:

$$-R = k(X) \quad (15.39)$$

After the system has operated for some time, it achieves a *steady state* such that:

$$\frac{dn_x}{dt} = 0 \quad (15.40)$$

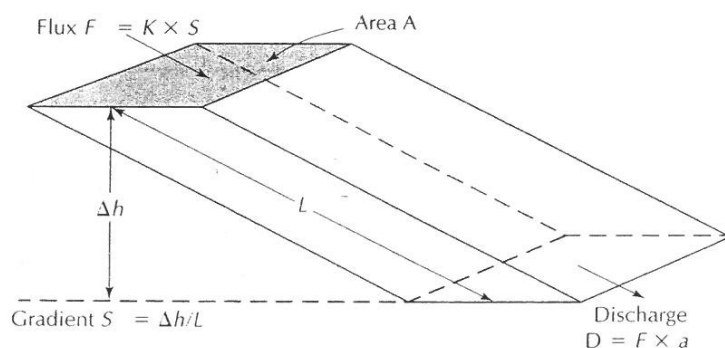


Figure 15.3 Schematic representation of a volume increment of a porous bed (aquifer) in subsurface. The movement of water through the aquifer is governed by Darcy's law expressed by equation 15.33 in terms of the flux F and the gradient S . The flux has units of mass (or volume) per unit area per unit time, which reduce to cm/sec. The flux F times the area a yields the discharge D , which is the volume of water leaving the system in unit time.

Under conditions of steady state and after substituting the rate equation for R , equation 15.38 becomes:

$$-k(X)dV = Dd(X) \quad (15.41)$$

which can be integrated between the inlet and outlet of the reactor (dV):

$$-\frac{k}{D} \int_0^V dV = \int_{(X)_1}^{(X)_2} \frac{d(X)}{(X)} \quad (15.42)$$

where $(X)_1$ and $(X)_2$ are the molar concentrations of X at the entrance and exit of volume dV , respectively. The result is:

$$-\frac{kV}{D} = \ln \frac{(X)_2}{(X)_1} \quad (15.43)$$

Note that V/D has the dimensions of time called the *contact time*:

$$\frac{V}{D} = \frac{\text{volume} \times \text{time}}{\text{volume}} = \text{time} \quad (15.44)$$

and therefore represents the average time molecule of X spends in the reactor. If we let $V/D = t$, equation 15.43 becomes:

$$\ln \frac{(X)_2}{(X)_1} = -kt \quad (15.45)$$

$$(X)_2 = (X)_1 e^{-kt} \quad (15.46)$$

Thus, equation 15.43 reduces to a first-order rate equation when it is applied to a static system.

15.3 Transport of Matter: Diffusion

Molecules or ions of a solute in a solution tend to disperse throughout a stationary solvent by a process known as *diffusion*. If the medium is itself in turbulent motion, then the dispersion of the solute is called *turbulent* or *eddy diffusion*. The flux (F) of ions or molecules diffusing through a stationary medium is proportional to the concentration gradient (dc/dx):

$$F = -D \left(\frac{dc}{dx} \right) \quad (15.47)$$

where D here represents the *diffusion coefficient* and the minus sign indicates that the concentration gradient has a negative slope; that is, the ions and molecules move from points of high concentration toward points of low concentration within the medium. Equation 15.47 was formulated by the German chemist Adolf Fick in the 1850s and is known as *Fick's first law*. The diffusion coefficient for ions and molecules in water are of the order of 10^{-5} cm²/sec, whereas for diffusion of ions in solids at temperatures around 1000°C, D has values around 10^{-10} cm²/sec. Diffusion coefficients for selected ions in dilute aqueous solutions are listed in Table 15.1.

In case a solute is transported both by advection and by diffusion, the total flux (F) of an ion or molecule X is the sum of the advective flux (F_a) and the diffusive flux (F_d) (Lerman, 1979, p. 58):

$$F = F_a + F_d \quad (15.48)$$

If the solvent moves with velocity v and the concentration of component X is (X) , the advective flux according to equation 15.32 is:

$$F_a = (X)v \quad (15.49)$$

Table 15.1 Diffusion Coefficients for Ions in Dilute Aqueous Solutions at 25°C

Cation	D , cm ² /sec	Anion	D , cm ² /sec
Li ⁺	10.3×10^{-6}	F ⁻	14.6×10^{-6}
Na ⁺	13.3×10^{-6}	Cl ⁻	20.3×10^{-6}
K ⁺	19.6×10^{-6}	Br ⁻	20.1×10^{-6}
Rb ⁺	20.6×10^{-6}	I ⁻	20.0×10^{-6}
Mg ²⁺	7.05×10^{-6}	HS ⁻	17.3×10^{-6}
Ca ²⁺	7.93×10^{-6}	HSO ₄ ⁻	13.3×10^{-6}
Sr ²⁺	7.94×10^{-6}	SO ₄ ²⁻	10.7×10^{-6}
Ba ²⁺	8.48×10^{-6}	NO ₃ ⁻	19.0×10^{-6}
Mn ²⁺	6.88×10^{-6}	HCO ₃ ⁻	11.8×10^{-6}
Fe ²⁺	7.19×10^{-6}	CO ₃ ²⁻	9.55×10^{-6}
Co ²⁺	6.99×10^{-6}	H ₂ PO ₄ ⁻	8.46×10^{-6}
Ni ²⁺	6.79×10^{-6}	HPO ₄ ²⁻	7.34×10^{-6}
Cu ²⁺	7.33×10^{-6}	PO ₄ ³⁻	6.12×10^{-6}
Zn ²⁺	7.15×10^{-6}	MoO ₄ ²⁻	9.91×10^{-6}

SOURCE: Lerman (1979), Table 3.1.

The diffusive flux is given by equation 15.47. Therefore, the total flux of component X is:

$$F_X = (X)v - D \frac{d(X)}{dx}$$

We may want to compare the fluxes caused by advection and by diffusion. For that purpose we define the advective transport distance (L_a) by the equation:

$$L_a = vt \quad (15.50)$$

Since the diffusion coefficient (D) has the dimensions of square centimeters per second, Dt has the dimensions of square centimeters and $(Dt)^{1/2}$ has dimensions of length only. Therefore, the diffusional transport distance (L_d) is:

$$L_d = (Dt)^{1/2} \quad (15.51)$$

and the ratio of transport distances is:

$$\frac{L_a}{L_d} = \frac{vt}{(Dt)^{1/2}} = v \left(\frac{t}{D} \right)^{1/2} \quad (15.52)$$

Evidently, for fixed values of v and D the ratio L_a/L_d increases with the square root of time. The time required to make the respective transport distances equal to each other is obtained by solving equation 15.52 for t in the case that $L_a/L_d = 1.00$:

$$v \left(\frac{t}{D} \right)^{1/2} = 1.00 \quad (15.53)$$

$$t = \frac{D}{v^2} \quad (15.54)$$

For example, let $D = 1 \times 10^{-5} \text{ cm}^2/\text{sec}$ and $v = 1.0 \text{ m/yr}$, which is fairly typical for the movement of groundwater through an aquifer. Converting the velocity to centimeters per second:

$$v = \frac{1(100)}{1(365.25)(24)(60)(60)} = 3.2 \times 10^{-6} \text{ cm/sec} \quad (15.55)$$

and from equation 15.54 for the case that $L_a = L_d$

$$\begin{aligned} t &= \frac{1 \times 10^{-5}}{(3.2 \times 10^{-6})^2} = 9.8 \times 10^5 \text{ sec} \\ &= 11.3 \text{ days} \end{aligned} \quad (15.56)$$

Evidently, after 11.3 days advection (based on a velocity of 1.0 m/yr) can transport the solute as far as diffusion could alone. Therefore, transport of solute in this case is primarily by advection rather than by diffusion. However, in a case where the velocity of the water is only $v = 1 \text{ cm/1000 yr}$, which is representative of the rate of sediment deposition in the oceans, then $v = 1/1000(365.25)(24)(60)(60) = 3.16 \times 10^{-11} \text{ cm/sec}$ and:

$$\begin{aligned} t &= \frac{1 \times 10^{-5}}{(3.16 \times 10^{-11})^2} = 1.00 \times 10^{16} \text{ sec} \\ &= 320 \times 10^6 \text{ yr} \end{aligned} \quad (15.57)$$

In this case, it takes a long time before transport by advection becomes equivalent to diffusion. This means that diffusion is the dominant transport mechanism in deep-sea sediment.

Fick's first law relates the flux of component X to the concentration gradient of X but does not indicate how the concentration of X varies with time. We therefore consider a volume increment having a unit cross section and a length dx extending from x to $x + dx$. The change in the concentration of X with time in this volume increment is the difference between the flux entering at x and leaving at $x + dx$, divided by dx . This is dimensionally correct because:

$$\begin{aligned} \frac{\text{flux}}{\text{length}} &= \frac{\text{g}}{(\text{cm}^2)(\text{cm})(\text{sec})} = \frac{\text{g}}{(\text{cm}^3)(\text{sec})} \\ &= \frac{\text{concentration}}{\text{time}} \end{aligned} \quad (15.58)$$

Therefore, based on Fick's first law (equation 15.47):

$$\frac{\partial c}{\partial t} = \frac{1}{dx} \left[-D \left(\frac{\partial c}{\partial x} \right)_x + D \left(\frac{\partial c}{\partial x} \right)_{x+dx} \right] \quad (15.59)$$

where the equation is written in terms of partial derivatives because we are differentiating only with respect to the x -direction while holding the other two directions, y and z , constant. The outgoing flux at $x + dx$ has a positive sign because it is being subtracted from the flux entering at x . Since:

$$\left(\frac{\partial c}{\partial x}\right)_{x+dx} = \left(\frac{\partial c}{\partial x}\right)_x + \frac{\partial}{\partial x} \left(\frac{\partial c}{\partial x}\right) dx \quad (15.60)$$

we substitute into equation 15.59 (Moore, 1955, p. 448) and obtain:

$$\begin{aligned} \frac{\partial c}{\partial t} &= -\frac{D}{dx} \left(\frac{\partial c}{\partial x}\right)_x + \frac{D}{dx} \\ &\times \left[\left(\frac{\partial c}{\partial x}\right)_x + \frac{\partial}{\partial x} \left(\frac{\partial c}{\partial x}\right) dx \right] \end{aligned} \quad (15.61)$$

which reduces to:

$$\frac{\partial c}{\partial t} = D \left(\frac{\partial^2 c}{\partial x^2}\right) \quad (15.62)$$

Equation 15.62 is *Fick's second law*, which states that the rate of change of the concentration of a component X is equal to the second derivative of the linear concentration gradient multiplied by the diffusion coefficient. Equation 15.62 has the same form as the equation that governs heat conduction and has been solved for many different boundary conditions (Crank, 1975).

The application of Fick's second law is illustrated by the case where a thin layer of the diffusing substance is placed between two layers of a stationary medium. According to Crank (1975), the solution of equation 15.62 is:

$$\frac{c}{c_0} = (\pi Dt)^{-1/2} e^{-x^2/4Dt} \quad (15.63)$$

When the concentration c is plotted as a function of distance x for different values of t using an appropriate value of the diffusion coefficient (D), a series of Gaussian error curves is obtained that show how the diffusing substance spreads out as time increases (Moore, 1955, p. 449). Three such curves are shown in Figure 15.4, where the diffusion takes place both in the $+x$ and in the $-x$ direction and the concentration must therefore be divided by 2 because diffusion occurs in two directions.

Equation 15.63 can be used to evaluate diffusion data. If we take natural logarithms of equation 15.63, we obtain:

$$\ln \left(\frac{c}{c_0}\right) = -\frac{1}{2} \ln(\pi Dt) - \frac{x^2}{4Dt} \quad (15.64)$$

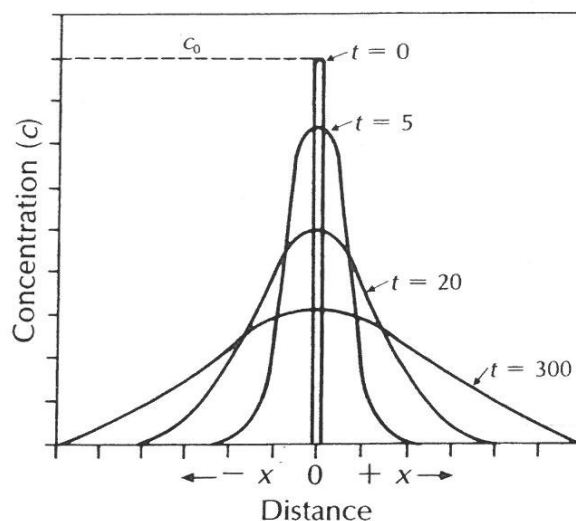


Figure 15.4 Schematic representation of the distribution of a diffusing substance that starts as a thin film in a stationary medium. The distribution of the diffusing substance at different times (t) is governed by Fick's second law expressed by equation 15.63.

which is a straight line in coordinates of $\ln(c/c_0)$ and x^2 . Therefore, measurements of the concentration of an ion in a body of water through which the ion is diffusing define a straight line in this coordinate system. The slope of this line is:

$$m = -\frac{1}{4Dt} \quad (15.65)$$

and its intercept on the y-axis is:

$$b = -\frac{1}{2} \ln(\pi Dt) \quad (15.66)$$

An example of diffusion of ions occurs in Lake Vanda in Wright Valley, southern Victoria Land, Antarctica. This lake contains a layer of dense brine from which ions have diffused into an overlying layer of less saline water. Green and Canfield (1984) explained the variation in the chemical composition of water taken at different depths in Lake Vanda in terms of diffusion of ions. The variation of the concentration of Ca^{2+} ions with depth is illustrated in Figure 15.5, which also contains a plot of $\ln \text{Ca}^{2+}/\text{Ca}_0^{2+}$ versus x^2 .

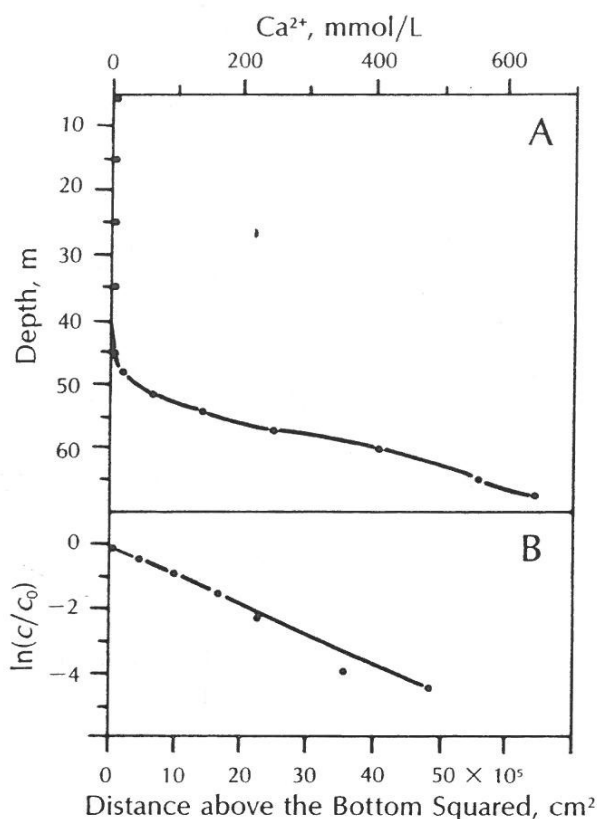


Figure 15.5 A: Variation of the concentration of Ca^{2+} as a function of depth in Lake Vanda, Wright Valley, southern Victoria Land, Antarctica. The increase of Ca^{2+} and other ions is caused by the presence of a layer of dense brine whose temperature is about 25°C , although Lake Vanda has a permanent ice cover that is 4 m thick (Green and Canfield, 1984).

B: Plot of Ca^{2+} concentrations in Lake Vanda in accordance with equation 15.64, which is based on Fick's second law of diffusion. The data points define a straight line whose slope was used in equation 15.68 to calculate an age of 1290 years for the dilute layer above the brine.

where x is the distance in centimeters from the bottom of the lake to each analyzed sample. The data points below a depth of 45 m in the lake fit a straight line whose slope is $m = -0.09096 \times 10^{-5}$. Since the water temperature *increases* with depth from $\sim 10^\circ\text{C}$ at 45 m to $\sim 25^\circ\text{C}$ at the bottom, the diffusion coefficient of Ca^{2+} was taken to be $D = 6.73 \times 10^{-6} \text{ cm}^2/\text{sec}$ at an average tempera-

ture of 18°C (Li and Gregory, 1974). We can now solve equation 15.65 for time in years. Since:

$$-\frac{1}{4Dt} = -0.09096 \times 10^{-5} \quad (15.67)$$

$$t = \frac{10.99 \times 10^5}{4(6.73 \times 10^{-6})(3.155 \times 10^7)} = 1290 \text{ yr} \quad (15.68)$$

This result indicates that the observed distribution of Ca^{2+} concentrations in Lake Vanda below a depth of ~ 45 m is attributable to ionic diffusion from the dense brine layer in the past 1290 years.

When diffusion involves the movement of charged particles, electrical neutrality throughout the system must be maintained either by the movement of positively and negatively charged species in the same direction or by the movement of charged particles in the opposite direction. Moreover, the magnitude of the diffusion coefficient of an ionic species is affected by electrostatic interactions with other ions depending on its ionic potential, that is, the charge-to-radius ratio (Lasaga, 1979). Therefore, the magnitudes of diffusion coefficients depend to some extent on the composition of the environment in which diffusion is taking place. In general, several different kinds of diffusion coefficients have been defined that apply to specific types of diffusion (Crank, 1975).

The magnitudes of diffusion coefficients are also strongly dependent on the temperature. The temperature dependence of D is expressed by the Arrhenius equation:

$$D = D_0 e^{-E_a/RT} \quad (15.69)$$

where D_0 is the *frequency factor*, E_a is the activation energy, R is the gas constant, and T is the temperature in kelvins. In this context, E_a is a measure of the energy required to start the movement of the diffusing substance or ion. We can linearize equation 15.69 by taking natural logarithms:

$$\ln D = \ln D_0 - \frac{E_a}{RT} \quad (15.70)$$

This is the equation of a straight line in coordinates of $\ln D$ and the reciprocal of the temperature whose slope is:

$$m = -\frac{E_a}{R} \quad (15.71)$$

Since the slope is negative, $\ln D$ increases with decreasing values of $1/T$, that is, with increasing temperature. The temperature dependence of diffusion coefficients is especially important in igneous petrology because it affects the distribution of cations in minerals that form at elevated temperatures by crystallization of silicate melts. Values of relevant diffusion coefficients of ions in melts and in minerals at 1000°C were compiled by Henderson (1982, Tables 8.1 and 8.2). Diffusion coefficients of ions and molecules through solids at 25°C are listed in Table 15.2.

15.4 Growth of Concretions During Diagenesis

Diagenesis includes all of the chemical, physical, and biological processes that take place in sediment after it was deposited. The *chemical* processes may include dissolution of mineral particles in the pore water, precipitation of insoluble compounds from the pore water, and ion exchange reactions between aqueous species and the surfaces of solids. *Biological* activity may stir up the sediment (bioturbation), and bacteria can bring about chemical changes, such as the reduction of sulfate to bisulfide ions. In addition, the presence of certain kinds of biogenic compounds may inhibit dissolution of mineral grains by coating their surfaces or enhance it by adsorbing some of the ions that are released. The chemical reactions and biological activity take place in an environment controlled by *physical* processes, including the deposition of sediment and its compaction as well as the flow of pore water in response to pressure gradients.

The description of these processes can be approached in three different ways all of which are useful in the study of diagenesis (Berner, 1980):

1. Observations regarding changes in the chemical composition of the pore water and the sediment with depth and hence with time.
2. Experimentation in the laboratory or thermodynamic calculations both of which may permit qualitative predictions about changes that should occur with depth as a result of diagenesis.
3. A combination of transport mechanisms (advection and diffusion) and measured rates of chemical reactions from which a transport-dependent reaction model of diagenesis can be developed.

Such a model was presented by Berner (1980) to describe the growth of a concretion during diagenesis.

Minerals that precipitate from the pore water of sediment may either form *within* the pore spaces as a cement or replace preexisting minerals that dissolve. Both processes may cause the formation of "localized mineral segregations" called *concretions* (Berner, 1980, p. 108). The reactants required to form a mineral cement may originate by dissolution of other minerals within the sediment by a process called *diagenetic redistribution*, or they may be transported to the site of deposition from distant sources. Similarly, the ions released by dissolution of the preexisting minerals may participate in the formation of the authigenic minerals at the same site (replacement), or they may be transported away.

After nucleation of an authigenic mineral has occurred, the continued growth of crystals depends on (1) the transport of ions to the surface of the crystal; (2) surface reactions such as adsorption, dehydration, and ion exchange; and (3) removal of the products of chemical reactions. The rate of growth of the crystals may be controlled by each of the three processes or by combinations of them. For the sake of simplicity, we consider only transport and surface chemistry as the rate-controlling processes of crystal growth.

In case crystal growth is controlled primarily by *transport mechanisms*, the concentrations of

Table 15.2 Diffusion Coefficients and Activation Energies of Ions in Solids at 25°C

<i>Matrix</i>	<i>Species</i>	<i>D, (25°C) cm²/sec</i>	<i>E_a kcal/mol</i>	<i>Direction or composition</i>
Quartz	Na	1.1×10^{-15}	20.2	parallel to c-axis
Quartz	Na	1.8×10^{-28}	41	90° to c-axis
Quartz	Ca	1.4×10^{-45}	68	
Na ₂ SiO ₃ glass	Na	8.2×10^{-17}	18.5	
K ₂ SiO ₃ glass	K	5.9×10^{-17}	17.5	
SiO ₂ glass	H ₂ O	3.4×10^{-12}	11.3	
Obsidian	H ₂ O	1.0×10^{-16}	13	H ₂ O = 0 wt. %
Obsidian	H ₂ O	8.0×10^{-16}	13	H ₂ O = 0.5 wt. %
Obsidian	H ₂ O	1.0×10^{-14}	13	H ₂ O = 6.0 wt. %
Feldspar	K	1.6×10^{-49}	68	orthoclase, Or ₉₄
Feldspar	Na	2.1×10^{-38}	53	orthoclase, Or ₉₄
Feldspar	⁴⁰ Ar	3.6×10^{-34}	43.1	orthoclase, Or ₉₄
Feldspar	O ₂	5.6×10^{-28}	29.6	microcline, Or ₁₀₀
Feldspar	O ₂	1.1×10^{-28}	29.6	adularia, Or ₁₀₀
Feldspar	O ₂	7.7×10^{-27}	25.6	adularia, Or ₉₈
Feldspar	O ₂	5.6×10^{-25}	21.3	albite, Ab ₉₇₋₉₉
Feldspar	O ₂	8.7×10^{-27}	26.2	anorthite, An ₉₆
Mica	Ar	1.7×10^{-43}	57.9	phlogopite, 4% annite, <i>P</i> = 2 kbar
Mica	Ar	1.4×10^{-22}	23.3	phlogopite in vacuum
Zeolite	H ₂ O	10^{-7} to 10^{-8}		heulandite, parallel to structure
Zeolite	H ₂ O	10^{-11}		heulandite, perpendicular to structure
Zeolite	H ₂ O	2.1×10^{-8}		Ca-heulandite
Zeolite	H ₂ O	1.3×10^{-7}		Ca-chabazite
Zeolite	H ₂ O	2.0×10^{-13}		Na-analcime
Zeolite	H ₂ O	1.5×10^{-12}		Na-analcime
Zeolite	Na	10^{-12} to 10^{-13}		analcime, chabazite
Zeolite	Cs	10^{-24} , 10^{-13}		analcime, chabazite
Zeolite	Ca	4×10^{-16}		chabazite
Calcite	CO ₂	1.4×10^{-46}	58.0	CaCO ₃ = 99.5% MgCO ₃ = 0.5% FeCO ₃ = trace

SOURCE: Lerman (1979), Table 3.9.

the reactants in the solution adjacent to the crystal are maintained at levels required for equilibrium and the crystals grow as rapidly as the reactants arrive at the surface of the crystal by advection or diffusion. If the rate of crystal growth is dominated by *surface reactions*, then

the concentrations of the reactants in the solution adjacent to the crystal may be in excess of equilibrium values and may remain uniform throughout the aqueous phase. In reality, *both* processes may affect the growth rate of crystals in sediment during diagenesis, but rapid growth of authigenic

crystals is generally associated with transport-controlled reactions.

The rate of dissolution of minerals is similarly controlled by the transport mechanism, by surface chemistry, or by a combination of the two. If dissolution of preexisting minerals and simultaneous precipitation of authigenic minerals are both transport controlled, then a zone of depletion develops around each growing crystal. In this process of diagenetic redistribution the thickness of the depleted zone around each crystal depends on the abundance of the dissolving mineral in the sediment and on the radius of the crystal as required by mass balance. If, on the other hand, the rate of dissolution and crystal growth are both entirely controlled by surface processes, then dissolution of preexisting minerals is selective—that is, smaller crystals dissolve faster than larger ones because they have more surface area—and growing crystals are surrounded only by a poorly defined zone of depletion whose thickness cannot be predicted by a mass-balance calculation.

Therefore, the presence of a well-defined zone of depletion in the sediment around a concretion indicates that it formed as a result of local diagenetic redistribution by a transport-dominated process. In this case, it is possible to determine how much time was required for the concretion to form.

We assume that a concretion is forming, as shown in Figure 15.6, by dissolution of minerals in the sediment at the outer edge of a spherical zone of depletion whose width is L and that the transport of reactants to the surface of the concretion is entirely by diffusion. Under these conditions Fick's first law (equation 15.47) applies, and the flux (F) of reactants arriving on the surface of the spherical concretion is (Berner, 1980; Nielsen, 1961; Frank 1950):

$$F_c = -\frac{\phi D}{L}(c_d - c_p) \quad (15.72)$$

where c is the molar concentration of the material in the pore water in equilibrium with the solid that is dissolving (d) or precipitating (p) in the zone of depletion and on the surface of the concretion, respectively. Therefore, $(c_d - c_p)/L$ is the concentration gradient across the zone of deple-

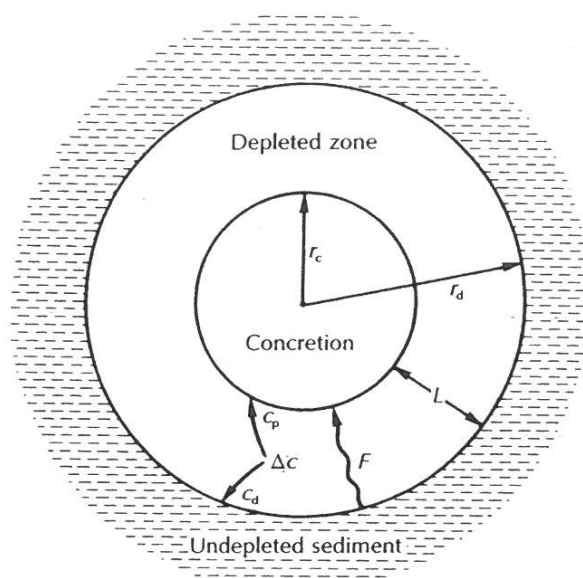


Figure 15.6 Growth of a concretion having a radius of r_c by diagenetic redistribution caused by dissolution of a mineral in the sediment, diffusion of the resulting ions to the surface of the concretion, and precipitation of a new mineral in the pore spaces of the sediment. The process is assumed to be transport dominated and therefore causes the formation of a depleted zone having a radius of r_d and a width L . In addition, V is the molar volume of the dissolving or precipitating minerals, ϕ is the porosity of the sediment, D is the diffusion coefficient, f is the volume fraction of the dissolving or precipitating mineral, and F is the flux of ions to the surface of the concretion and leaving the outer surface of the depleted zone. The model discussed here was presented by Berner (1980).

tion. We can simplify equation 15.72 by replacing $c_d - c_p$ by Δc :

$$F_c = -\frac{\phi D \Delta c}{L} \quad (15.73)$$

The flux due to diffusion is diminished by a factor ϕ , which is the *porosity* of the sediment defined as the fraction of a unit volume of sediment that consists of water-filled pore spaces.

The concretion grows as a function of time by the addition of a certain amount of mass (dM_p) to its surface:

$$\frac{dM_p}{dt} = -4\pi(r_c)^2 F_c \quad (15.74)$$

where $4\pi r_c^2$ is the surface area of the spherical concretion of radius r_c and F_c is the flux of material being delivered to the surface of the concretion by diffusion alone. Equation 15.74 is dimensionally correct because:

$$4\pi(r_c)^2 F_c = \frac{\text{length}^2 \times \text{mass}}{\text{length}^2 \times \text{time}} = \frac{\text{mass}}{\text{time}} \quad (15.75)$$

In addition, the minus sign is required because dM/dt must remain positive when equation 15.73 is substituted for F_c in equation 15.74.

The mass increment (dM) added to the concretion is also given by:

$$dM_p = \frac{4\pi(r_c)^2}{V_p} f_p dr_c \quad (15.76)$$

where f_p is the volume fraction of the precipitating material in the concretion and V_p is the molar volume of the precipitation material. Molar volumes of common minerals are listed in Table 15.3. The dimensions of equation 15.76 are in terms of mass because f_p is a dimensionless number and:

$$\frac{4\pi}{V_p} (r_c)^2 f_p dr_c = \frac{\text{length}^2 \times \text{length}}{\text{length}^3/\text{mass}} = \text{mass} \quad (15.77)$$

We now substitute equation 15.73 for F_c and equation 15.76 for dM_p into equation 15.74. The result is:

$$\frac{4\pi(r_c)^2}{V_p} f_p \left(\frac{dr_c}{dt} \right) = 4\pi(r_c)^2 \phi D \left(\frac{\Delta c}{L} \right) \quad (15.78)$$

which reduces to:

$$\frac{dr_c}{dt} = \frac{\phi V_p D}{f_p} \left(\frac{\Delta c}{L} \right) \quad (15.79)$$

Unfortunately, we cannot integrate this equation because the width (L) of the depleted zone also increases with time as the concretion grows. Therefore, we must express L as a function of time before we can use equation 15.79 to derive a relationship between the radius of a concretion and the time required for it to grow to a given size.

The thickness (L) of the zone of depletion increases with time at a rate that depends on the rate of growth of the concretion:

Table 15.3 Molar Volumes of Common Minerals in Sedimentary Rocks

Mineral	Formula	Molar volume, cm^3
Troilite (hex.)	FeS	18.20
Pyrrhotite (hex.)	$\text{Fe}_{0.98}\text{S}$	18.11
Alabandite (cubic)	MnS	21.457
Pyrite (cubic)	FeS_2	23.940
Marcasite (orth.)	FeS_2	24.579
Boehmite (orth.)	$\text{AlO}(\text{OH})$	19.535
Diaspore (orth.)	$\text{AlO}(\text{OH})$	17.760
Gibbsite (mon.)	$\text{Al}(\text{OH})_3$	31.956
Hematite (hex.)	Fe_2O_3	30.274
Magnetite (cubic)	Fe_3O_4	44.524
Goethite (orth.)	$\alpha\text{-FeO}(\text{OH})$	20.82
Manganosite (cubic)	MnO	13.221
Pyrolusite (tet.)	MnO_2	16.61
Bixbyite (cubic)	Mn_2O_3	31.37
Hausmannite (tet.)	Mn_3O_4	46.95
Quartz (α) (hex.)	SiO_2	22.688
Quartz (β) (hex.)	SiO_2	23.718
Cristobalite (α) (tet.)	SiO_2	25.739
Cristobalite (β) (cubic)	SiO_2	27.381
Tridymite (β) (hex.)	SiO_2	27.414
Uraninite (cubic)	UO_2	24.618
Calcite (hex.)	CaCO_3	36.934
Aragonite (orth.)	CaCO_3	34.15
Siderite (hex.)	FeCO_3	29.378
Magnesite (hex.)	MgCO_3	28.018
Rhodochrosite (hex.)	MnCO_3	31.073
Dolomite (hex.)	$\text{MgCa}(\text{CO}_3)_2$	64.341
Barite (orth.)	BaSO_4	52.10
Anhydrite (orth.)	CaSO_4	45.94
Gypsum (mon.)	$\text{CaSO}_4 \cdot 2\text{H}_2\text{O}$	74.69

SOURCE: Weast et al. (1986).

$$\frac{dL}{dt} = \frac{dr_d}{dt} - \frac{dr_c}{dt} \quad (15.80)$$

where the subscripts identify the zone of depletion (d) and the concretion (c) as before and r_d is

the radius of the depleted zone. The mass of the dissolving material is:

$$\frac{dM_d}{dt} = 4\pi(r_d)^2 F_d \quad (15.81)$$

The rate of dissolution (dM_d/dt) in equation 15.81 is a negative quantity because the material is lost from the system. Therefore, the two minus signs associated with F and with the loss of material cancel each other.

The amount of material lost from the zone of depletion is also expressible by an equation that is similar to 15.76:

$$dM_d = -\frac{4\pi}{V_d}(r_d)^2 f_d dr_d \quad (15.82)$$

where f_d is the volume fraction of the dissolving mineral in the sediment and V_d is the molar volume of the dissolving material. We can now substitute equation 15.82 for dM_d and equation 15.73 for F_d in equation 15.81. The result is:

$$-\frac{4\pi}{V_d}(r_d)^2 f_d \left(\frac{dr_d}{dt}\right) = -4\pi(r_d)^2 \phi D \left(\frac{\Delta c}{L}\right) \quad (15.83)$$

which reduces to:

$$\frac{dr_d}{dt} = \frac{\phi V_d D}{f_d} \left(\frac{\Delta c}{L}\right) \quad (15.84)$$

Next, we substitute the rate of growth of the concretion (equation 15.79) and the rate of growth of the depleted zone (equation 15.84) into equation 15.80, which expresses the rate of change of the width (L) of the depleted zone:

$$\frac{dL}{dt} = \frac{\phi V_d D}{f_d} \left(\frac{\Delta c}{L}\right) - \frac{\phi V_p D}{f_p} \left(\frac{\Delta c}{L}\right) \quad (15.85)$$

which yields:

$$\frac{dL}{dt} = \phi D \left(\frac{\Delta c}{L}\right) \left(\frac{V_d}{f_d} - \frac{V_p}{f_p}\right) \quad (15.86)$$

Equation 15.85 can be integrated in the form:

$$\int L dL = \text{constant} \int dt \quad (15.87)$$

If $L = 0$ when $t = 0$, the result is $\frac{1}{2}L^2 = \text{constant} \times t$ from which it follows that:

$$L = \left[2\phi D \Delta c \left(\frac{V_d}{f_d} - \frac{V_p}{f_p} \right) t \right]^{1/2} \quad (15.88)$$

We are now able to integrate equation 15.79 and thereby obtain the desired relationship between the radius of the growing concretion (r_d) and time by substituting equation 15.88 into equation 15.79. In order to simplify the algebra we write equation 15.88 as:

$$L = G^{1/2} t^{1/2} \quad (15.89)$$

and equation 15.79 as:

$$\frac{dr_c}{dt} = \frac{H}{L} \quad (15.90)$$

where G and H are constants defined by equations 15.88 and 15.79, respectively. Therefore:

$$dr_c = \left(\frac{H}{G^{1/2}} \right) t^{-1/2} dt \quad (15.91)$$

By integrating equation 15.91 for the condition that $r_c = 0$ when $t = 0$, we obtain:

$$r_c = \left(\frac{2H}{G^{1/2}} \right) t^{1/2} \quad (15.92)$$

or

$$t = \left(\frac{r_c}{2} \right)^2 \frac{G}{H^2} \quad (15.93)$$

The ratio of the constant G/H^2 can now be simplified:

$$\frac{G}{H^2} = \frac{2\phi D \Delta c (V_d/f_d - V_p/f_p) f_p^2}{\phi^2 V_p^2 D^2 \Delta c^2} \quad (15.94)$$

$$\begin{aligned} \frac{G}{H^2} &= \frac{2(V_d/f_d - V_p/f_p) f_p^2}{\phi V_p^2 D \Delta c} \\ &= \frac{2(V_d f_p - V_p f_d) f_p^2}{\phi V_p^2 D \Delta c f_p} \end{aligned} \quad (15.95)$$

Next, we assume that the precipitating mineral fills the available pore spaces in the sediment ($f_p = \phi$) and that the molar volumes of the dissolving and precipitating minerals are approximately equal ($V_d \approx V_p$):

$$\frac{G}{H^2} = \frac{2(\phi - f_d)}{V_p F_d D \Delta c} \quad (15.96)$$

Therefore, equation 15.93 becomes:

$$t = \frac{r_c^2(\phi - f_d)}{2V_p f_d D \Delta c} = \frac{r_c^2}{2V_p D \Delta c} \left(\frac{\phi}{f_d} - 1 \right) \quad (15.97)$$

This equation, derived by Berner (1980), is applicable to the growth of calcite concretions by dissolution of aragonite crystals disseminated throughout a body of sediment. If transport of ions is by diffusion and if $D = 2 \times 10^{-6} \text{ cm}^2/\text{sec}$, $\phi = f_p = 0.60$, and V (calcite) $\approx V$ (aragonite) = $35 \text{ cm}^3/\text{mol}$, then equation 15.97 can be solved for t for selected values of the abundance of aragonite crystals in the sediment (f_d). The concentration gradient (Δc) is the difference in the equilibrium concentrations of Ca^{2+} or CO_3^{2-} ions at the surfaces of aragonite and calcite crystals. At 25°C this difference is about $1.0 \times 10^{-5} \text{ mol/kg}$ or $1 \times 10^{-8} \text{ mol/g}$, based on the solubility product constants in Table 10.1. Therefore, if the aragonite content of the sediment is 15% ($f_d = 0.15$), a calcite concretion having a radius of 5.0 cm will form in:

$$\begin{aligned} t &= \frac{(5.0)^2(0.60/0.15 - 1)}{2 \times 35 \times 2 \times 10^{-6} \times 1 \times 10^{-8}} \\ &= 0.5357 \times 10^{14} \text{ sec} \end{aligned} \quad (15.98)$$

Since $1 \text{ yr} = 3.155 \times 10^7 \text{ sec}$,

$$t = \frac{0.5357 \times 10^{14}}{3.155 \times 10^7} = 1.7 \times 10^6 \text{ yr} \quad (15.99)$$

We see from this result that a calcite concretion the size of a small cannonball can grow in ~ 1.7 million years, which is quite fast on a geological time scale. In general, the *rate of growth* of concretions increases with the abundance of the dissolving mineral in the sediment. The relationship between the radius of a concretion and time, expressed by equation 15.97, is shown graphically in Figure 15.7 for the case where the aragonite content of the sediment is 5, 10, and 20% by volume (i.e., $f_d = 0.05, 0.10$, and 0.20). The resulting curves can be used to determine the time when a particular layer within a concretion was deposited. Therefore, we can regard carbonate concretions as *geochemical recorders* that have preserved a record of changes in the environment in which they formed. Concretions can reveal not only the

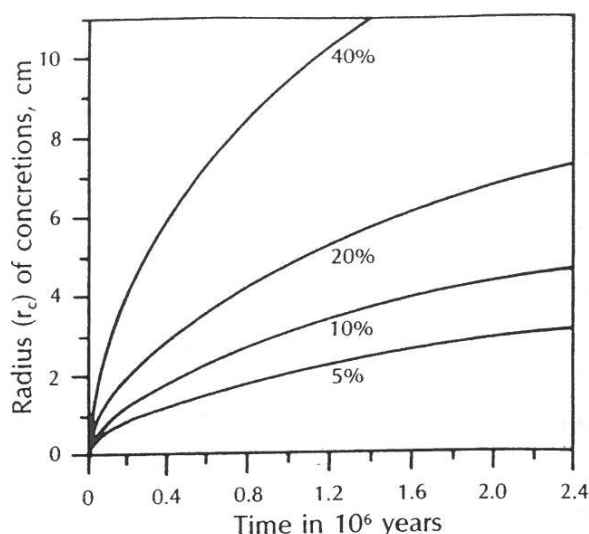


Figure 15.7 Growth of calcite concretions by dissolution of aragonite during diagenesis of sediment containing varying initial concentrations of aragonite (see equation 15.97 after Berner, 1980). Transport of ions is by diffusion, the concentration gradient across the zone of depletion is $1 \times 10^{-5} \text{ mol/kg}$, the diffusion coefficient $D = 2 \times 10^{-6} \text{ cm}^2/\text{sec}$, the porosity of the sediment in the depleted zone is 60%, which is assumed to be equal to the abundance of calcite in the concretions, and the molar volumes of calcite and aragonite are both taken to be $35 \text{ cm}^3/\text{mol}$. The curves indicate that the size of calcite concretions that grow in unit time increases with increasing abundance of aragonite in the sediment.

sequence of geochemical changes but also the length of the time intervals between them based on equation 15.97. Since concretions form soon after deposition of the sediment, they can provide a time scale for geochemical events that may have affected their mineral compositions, their growth rate, trace element concentrations, and the isotope compositions of O, C, Sr, Pb, and Nd they contain.

15.5 Growth of Monomineralic Layers

The sediment that is accumulating in deep-sea basins and in some lakes becomes anoxic because O_2 is consumed by organisms and by the oxidation

of organic matter. As a result, the solubility of the oxides in Fe and Mn increases as the environmental Eh decreases, as shown in Figures 14.6 and 14.7B. (The stability of Mn oxides in terms of Eh and pH is the subject of Problems 7–9 in Chapter 14.) Therefore, a concentration gradient of Fe^{2+} and Mn^{2+} ions may develop, which causes these ions to diffuse upward toward the sediment–water interface. When the ions return to the more oxygenated environment, they precipitate as oxyhydroxides of Fe^{3+} and Mn^{4+} because of the decrease in the solubilities of these compounds at the higher Eh (and pH) conditions in the water above the sediment. As a result, a layer of Fe and Mn oxides and hydroxides accumulates at the sediment–water interface, provided that the process is not terminated prematurely by burial caused by rapid sedimentation. The dissolution of scattered particles of Fe- and Mn-bearing minerals and the precipitation of authigenic oxides at the sediment–water interface is an example of diffusion-controlled diagenetic redistribution. The rate of growth of the monomineralic layer at the sediment–water interface can therefore be expressed by the equations derived in the preceding section for the growth of concretions.

The flux (F) at the sediment–water interface is given by Fick's first law (Berner, 1980):

$$F = -\phi D \frac{c_d - c_p}{L} \quad (15.100)$$

where c_d and c_p are the concentrations of Fe^{2+} and Mn^{2+} in the water in equilibrium with the dissolving mineral grains (c_d) and in equilibrium with the authigenic minerals at the sediment–water interface (c_p). Since L is the thickness of the depleted zone shown in Figure 15.8 ($(c_d - c_p)/L$ is the concentration gradient across the depleted zone. In addition, ϕ is the porosity and D is the diffusion coefficient defined as in Section 15.4.

In order to constrain the model to the simplest set of circumstances we specify that:

1. The authigenic minerals form within the oxygenated water on top of the sediment.
2. Transport of ions by advection is negligible.
3. The ions are not adsorbed on the surfaces of particles in the sediment.

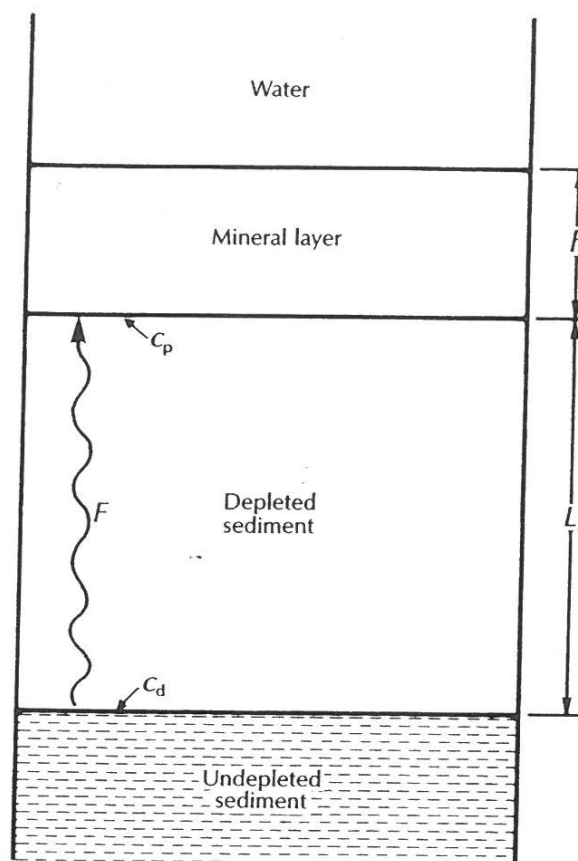


Figure 15.8 Formation of a layer of minerals at the sediment–water interface by diffusion of ions such as Fe^{2+} and Mn^{2+} from the sediment. The ions form by dissolution of particles in the sediment, which is anoxic compared to the water above; that is, it has a lower Eh. The ions precipitate when they reach the sediment–water interface to form a layer whose thickness (h) increases with time in accordance with equation 15.110. The thickness of the depleted layer (L) also increases with time (equation 15.104). However, calculations show that a layer of Fe or Mn oxyhydroxides can accumulate only when the rate of deposition of detrital sediment is virtually zero.

4. Sediment deposition from above is negligible.

By analogy with equation 15.79, the rate of growth of the monomineralic layer of thickness h is:

$$\frac{dh}{dt} = \frac{\phi V_p D}{f_p} \left(\frac{\Delta c}{L} \right) \quad (15.101)$$

where V_p is the molar volume of the precipitating phase and f_p is its volume fraction in the layer. Similarly, the rate of growth of the zone of depletion is given by equation 15.84 because the monomineralic layer is not expanding into the zone of depletion but is growing away from it. Therefore:

$$\frac{dL}{dt} = \frac{\phi V_d D}{f_d} \left(\frac{\Delta c}{L} \right) \quad (15.102)$$

Integration of this equation for the case that $L = 0$ when $t = 0$ yields:

$$\frac{1}{2} L^2 = \left(\frac{\phi V_d D \Delta c}{f_d} \right) t \quad (15.103)$$

and therefore:

$$L = \left(\frac{2\phi V_d D \Delta c}{f_d} \right)^{1/2} t^{1/2} \quad (15.104)$$

By substituting equation 15.104 in the form $L = G'^{1/2} t^{1/2}$ into equation 15.101 in the form $dh/dt = H'/L$, we obtain:

$$\frac{dh}{dt} = \frac{H' t^{-1/2}}{G'^{1/2}} \quad (15.105)$$

which we integrate for the condition that $h = 0$ when $t = 0$. The result is:

$$h = \frac{2H'}{G'^{1/2}} t^{1/2} \quad (15.106)$$

or

$$t = \left(\frac{h}{2} \right)^2 \frac{G'}{H'^2} \quad (15.107)$$

which is similar to equation 15.93. Next, we evaluate the ratio of the constant terms:

$$\frac{G'}{H'^2} = \frac{2\phi V_d D \Delta c f_p^2}{f_d \phi^2 V_p^2 D^2 (\Delta c)^2} = \frac{2V_d f_p^2}{f_d \phi V_p^2 D \Delta c} \quad (15.108)$$

If $V_p \approx V_d$, we obtain:

$$\frac{G'}{H'^2} = \frac{2f_p^2}{f_d \phi V_p D \Delta c} \quad (15.109)$$

and therefore equation 15.107 becomes:

$$t = \frac{h^2 f_p^2}{2f_d \phi V_p D \Delta c} \quad (15.110)$$

Note that in this case the volume fraction of the precipitating phase (f_p) is not equal to the porosity (ϕ) of the underlying sediment because the layer is forming in the water above the sediment.

Berner (1980) evaluated equation 15.110 for the case that $h = 1$ cm, $f_p = 0.2$ (flocculant precipitate), $\phi = 0.8$, $V_p \approx V_d = 20$ cm³/mol, $D = 3 \times 10^{-6}$ cm²/sec or 100 cm²/yr. According to equation 15.110, the time required to form a monomineralic layer of specified thickness increases with *decreasing* abundance (f_d) of the dissolving mineral in the sediment and with *decreasing* concentration gradient (Δc). If $f_d = 0.002$ (0.2% by volume) and $\Delta c = 10^{-4}$ mol/kg, equation 15.110 yields:

$$\begin{aligned} t &= \frac{(1)^2 (0.2)^2 (10^3)}{2(0.002)(0.8)(20)(100)(10^{-4})} \\ &= \frac{40 \times 10^4}{6.4} = 62,500 \text{ yr} \end{aligned} \quad (15.111)$$

The time required to form such a monomineralic layer is apparently quite *long* compared to normal sedimentation rates in the oceans or in lakes. Even on the abyssal plains of the oceans the sedimentation rate is of the order of 0.1 cm/1000 yr. Therefore, the layer of *detrital* sediment deposited in 62,500 yr is more than six times thicker than the 1-cm monomineralic layer; thus such a layer cannot form under normal circumstances in the oceans or in lakes. The existence of such monomineralic layers of Fe and Mn oxides in sediment therefore requires either that a period of sediment starvation occurred at the site, or that the upward diffusive flux of ions was augmented by an advective flux caused by the movement of water, or both.

The mathematical approach to the study of geochemical processes involving transport of reactants in different kinds of environments on the surface of the Earth is presented in the textbooks by Berner (1971, 1980) and by Lerman (1979). These books combine elegant mathematical descriptions of complex geological processes with analytical measurements and with understanding based on thermodynamics. In this way we can actually achieve the high level of *under-*

standing of geochemical processes, which is the principal objective of the science of geochemistry. The grandeur of the processes that convert sediment at the bottom of the oceans into sedimentary rocks was appreciated more than 100 years ago by the geologist K. von Gumbel (1888), who wrote (freely translated):

There is no doubt that magnificent chemical processes are taking place on the bottom of the ocean. They explain the partial reconstitution (diagenesis) and solidification of sediment by the formation of cements as exemplified by the marine deposits of the geologic past. (From a quotation in Berner, 1980.)

15.6 Summary

Geochemical processes, such as weathering or diagenesis, consist of sequences of chemical reactions, which are commonly dependent on the transport of reactants and products. The chemical reactions themselves involve complex mechanisms, each of which takes place at a characteristic rate. As a result, the rates of geochemical processes depend both on the rates of the chemical reactions and on the rates and methods of transport that contribute reactants and remove products from the reaction site.

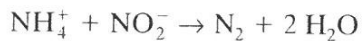
Geochemical processes can be studied by three methods. The first consists of analyzing appropriately chosen samples representing different stages or aspects of the process. Second, predictions about the outcome of geochemical processes can be made from thermodynamic considerations using various graphical representations or computations. In the third method, the entire process may be expressed in the form of equations that combine information about reaction mechanisms and their rates with appropriate transport models.

All three methods of study need to be employed because without analytical data and without a sense of direction provided by thermodynamics, mathematical modeling of geochemical processes cannot succeed. However, when done correctly, mathematical representations of geochemical processes achieve the highest level of understanding because they provide not only an explanation of what has already happened, but also permit predictions of what may happen in the future.

Geochemistry acquires a special importance among the sciences because it helps us to understand the world we live in and to make use of the resources it contains. The understanding we have of the environment we live in today permits us to anticipate how it may respond in the future to various anthropogenic perturbations.

Problems

1. Determine the *order* and write the rate equation of the reaction between ammonium ion and nitrite ion in aqueous solution at 30°C based on the following information about the initial rates of this reaction (Brown, 1968).



The concentration of NH_4^+ is kept constant so that the initial rates of the reaction depend only on the concentration of the nitrite ion.

Concentration of NO_2^- mol/L	Initial rate, mol L ⁻¹ sec ⁻¹
2.4×10^{-3}	18×10^{-8}
4.9×10^{-3}	33×10^{-8}
10.0×10^{-3}	65×10^{-8}
24.9×10^{-3}	156×10^{-8}
50.7×10^{-3}	338×10^{-8}
94.0×10^{-3}	643×10^{-8}

(Answer: $n \approx 1$)

2. The decomposition of gaseous nitrogen pentoxide (N_2O_5) is a first-order reaction where the concentration of N_2O_5 remaining is measurable by its pressure (Moore, 1955). Given the data below, determine the rate constant of this reaction.

Pressure of N_2O_5 , mm	Time, sec
348.4	0
247	600
185	1200
140	1800
105	2400
78	3000
58	3600
44	4200
33	4800
24	5400
18	6000
10	7200
5	8400
3	9600

3. Calculate the reaction half-life of the decomposition of N_2O_5 based on the data in Problem 2.

4. The rate constants for the formation of hydrogen iodide ($\text{H}_2 + \text{I}_2 \rightarrow 2 \text{HI}$) vary with temperature in the following manner (Moore, 1955).

k , sec^{-1}	T , $^\circ\text{C}$
316.2	473
100	441
31.6	412
14.1	394
3.16	362
2.20	352
0.316	320
0.251	315
0.0354	282

Determine the activation energy (E_a) of this reaction.

5. A major artesian aquifer in Florida has a vertical drop of 10 m over a map distance of 100 km. The hydraulic conductivity of this aquifer is $K = 10^{-1} \text{ cm/sec}$ (Lerman, 1979).

(a) Calculate the rate of flow of water in this aquifer in centimeters per year.

(b) Calculate the residence time of water in this aquifer, that is, the time required for a "particle" of water to traverse the entire aquifer.

6. The chloride concentrations of Lake Vanda, Wright Valley, southern Victoria Land, Antarctica, varies with depth as shown below (Green and Canfield, 1984).

Depth, m	Cl^- , mmol/L
5	8.26
15	10.04
25	13.91
35	14.39
45	29.34
48	43.72
51	231.88
54	496.5
57	846.26
60	1390.7
65	1900.2
67	2268.0

Interpret these data as in Figure 15.5 to determine the age of the dilute layer, assuming that $D = 17.1 \times 10^{-6} \text{ cm}^2/\text{sec}$ for Cl^- at 18°C .

7. Calculate the time required for a calcite concretion to grow by dissolution of disseminated aragonite and diffusion of the resulting Ca^{2+} ions at 10°C , based on the following information. The radius of the concretion is 8.0 cm, the diffusion coefficient of Ca^{2+} is $5.4 \times 10^{-6} \text{ cm}^2/\text{sec}$, the porosity of the sediment is 0.75, the volume fraction of aragonite in the sediment is 0.15, the difference in equilibrium concentration of Ca^{2+} in the pore water on the surfaces of calcite and aragonite crystals is $2.0 \times 10^{-5} \text{ mol/kg}$, the molar volume of calcite is $36.934 \text{ cm}^3/\text{mol}$, and the volume fraction of the authigenic calcite is equal to the porosity of the sediment.

(Answer: $1.0 \times 10^6 \text{ yr}$)

8. What is the width of the depleted zone around the concretion described in Problem 7? Obtain the molar volume of aragonite from Table 15.3.

(Answer: 30 cm)

9. Calculate the time required for a siderite concretion to grow and the width of the depleted zone around it given the following information: $r_c = 50 \text{ cm}$, $\Delta c = 2.4 \times 10^{-7} \text{ mol/g}$, $\phi = f_p = 0.85$, $D(\text{Fe}^{2+}) = 7.2 \times 10^{-6} \text{ cm}^2/\text{sec}$, $V_d \approx V_p = 30 \text{ cm}^3/\text{mol}$, $f_d = 0.30$.

(Answer: $t = 2.5 \times 10^6 \text{ yr}$, $L = 120 \text{ cm}$)

References

- BERNER, R. A., 1971. *Principles of Chemical Sedimentology*. McGraw-Hill, New York, 240 pp.
- BERNER, R. A., 1980. *Early Diagenesis. A Theoretical Approach*. Princeton University Press, Princeton, NJ, 241 pp.
- BROWN, T. L., 1968. *General Chemistry*, 2nd ed. Merrill, Columbus, OH, 668 pp.
- CRANK, J., 1975. *The Mathematics of Diffusion*, 2nd ed. Oxford University Press, London, 414 pp.
- FRANK, F. C., 1950. Radial symmetric growth controlled by diffusion. *Proc. Roy. Soc. London*, **A-201**:48-54.
- GREEN, W. J., and D. E. CANFIELD, 1984. Geochemistry of the Onyx River (Wright Valley, Antarctica) and its role in the chemical evolution of Lake Vanda. *Geochim. Cosmochim. Acta*, **48**(12):2457-2468.
- HENDERSON, P., 1982. *Inorganic Geochemistry*. Pergamon Press, Oxford, England, 353 pp.
- HOFMANN, A. W., B. J. GILETTI, H. S. YODER, JR., and R. A. YUND, 1974. *Geochemical Transport and Kinetics*. Carnegie Institution of Washington, Pub. 634, 353 pp.
- LASAGA, A. C., 1979. The treatment of multicomponent diffusion and ion pairs in diagenetic fluxes. *Amer. J. Sci.*, **279**:324-346.
- LASAGA, A. C., and R. J. KIRKPATRICK (Eds.), 1981. Kinetics of geochemical processes. *Rev. Mineral.*, **8**, Mineral. Soc. Amer., Washington, DC, 398 pp.
- LERMAN, A., 1979. *Geochemical Processes. Water and Sediment Environments*. Wiley, New York, 481 pp.
- LI, Y. -H., and S. GREGORY, 1974. Diffusion of ions in sea water and in deep-sea sediments. *Geochim. Cosmochim. Acta*, **38**: 703-714.
- MOORE, W. J., 1955. *Physical Chemistry*, 2nd ed. Prentice-Hall, Upper Saddle River, NJ, 633 pp.
- NIELSEN, A. E., 1961. Diffusion-controlled growth of a moving sphere. The kinetics of crystal growth in potassium perchlorate precipitation. *J. Phys. Chem.*, **65**:46-49.
- WEAST, R. C., M. J. ASTLE, and W. H. BEYER (Eds.), 1986. *CRC Handbook of Chemistry and Physics*. CRC Press, Boca Raton, FL.

IV

ISOTOPE GEOCHEMISTRY AND MIXING

The decay of long-lived radioactive isotopes of certain elements generates heat in the crust and mantle of the Earth and enables us to measure the ages of rocks and minerals. In addition, the fractionation of the stable isotopes of certain elements of low atomic number provides useful information about processes occurring in the hydrosphere, biosphere, and lithosphere. The isotope compositions of certain elements and the chemical compositions of water and rocks are also being altered by mixing of two, three, or more components. The chapters in Part IV deal with all of these subjects because they are important aspects of geochemistry.

Isotopic Geochronometers

The discovery of the radioactivity of uranium salts by Henry Becquerel in 1896 started a revolution in physics. Important advances soon resulted from the pioneering work of Marie Skłodowska-Curie and her husband Pierre in Paris and by the studies of Ernest Rutherford and Frederick Soddy in Montreal. By 1902 radioactivity was recognized as a process associated with several elements, some of which were newly discovered. It was also recognized that the process results in the emission of alpha and beta particles and of gamma radiation. The rate of emission is proportional to the number of radioactive atoms remaining and therefore decreases exponentially with time. These are the roots of atomic physics that, in the course of the 20th century, led to nuclear physics and to the inquiry into the most primitive forms of matter (Romer, 1971).

The importance of radioactivity to geology became apparent after Curie and Laborde (1903) reported that radioactivity is an exothermic process. This fact invalidated Lord Kelvin's calculations of the age of the Earth, which were based on the assumption that the Earth had cooled from its original molten state and does not have a source of heat of its own (Burchfield, 1975). By 1905 Rutherford was using the amount of He (alpha particles) in U ore to calculate geologic ages of the order of 500 million years, which were far in excess of Lord Kelvin's estimates of the age of the Earth (Rutherford, 1906; Thomson, 1899). The rest, as they say, is history. In 1909 John Joly summarized his measurements of the

radioactivity of rocks in a small book in which he also calculated the resulting rate of heat production. In 1911 Arthur Holmes published the first of his many papers on the dating of rocks based on the decay of U to Pb. In 1913 he published a book entitled *The Age of the Earth*, which contained the first geologic time scale based on age determinations of rocks and minerals.

These are the beginnings of a new discipline in the Earth Sciences called *Isotope Geology*, which seeks to use natural variations in the abundances of the isotopes of certain elements to study geological processes. The growth of this discipline has been closely linked to the design and subsequent improvement of analytical equipment, such as mass spectrometers and sensitive detectors of ionizing radiation. The entire subject has been presented in a textbook by Faure (1986). In addition, many earth science journals now publish papers based on or containing information derived from isotopic data. For this reason, we will briefly review the principles of isotope geology so that we can refer to them later as we take up different topics in geochemistry.

16.1 Decay Modes

The nucleosynthesis process in the ancestral stars (Chapter 2) produced more than 2200 nuclides, most of which were unstable and decomposed spontaneously in a short time. Only those nuclides that are stable or decay slowly have survived to the present. In addition, certain short-

lived radioactive nuclides (*radionuclides*) exist naturally because:

1. They are the short-lived daughters of U and Th, both of which consist of long-lived radioactive isotopes.
2. They are produced by nuclear reactions between cosmic rays and atoms in the atmosphere, as well as in rocks exposed at the surface.

A third group of radionuclides now exists on the Earth because of nuclear reactions caused by humans. The reactions consist primarily of induced fission of ^{235}U in nuclear reactors, but also include the detonation of fission and fusion devices in the atmosphere and in the subsurface. The anthropogenic radioactive contaminants produced by these activities include a very large number of radionuclides: ^3_1H (tritium), $^{14}_6\text{C}$, $^{90}_{38}\text{Sr}$, $^{129}_{53}\text{I}$, $^{137}_{55}\text{Cs}$, and many others. The containment of this radioactive waste has become a serious problem in environmental geochemistry to be discussed in Chapter 24.

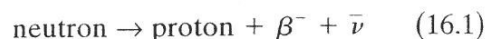
The large number of nuclides, whether stable or unstable, can be accommodated on a diagram called the *chart of the nuclides* (Chapter 2). This diagram is a plot of the atomic number (Z) versus the neutron number (N) and consists of squares, each of which is identified by the values of its N and Z coordinates. Figure 16.1 is such a chart for elements of low atomic number starting with hydrogen. The stable isotopes of each element are *shaded* and form a band that extends diagonally across the diagram. The short-lived radioactive isotopes are *unshaded* and form horizontal rows on either side of the stable isotopes. The ratio of neutrons to protons of the stable isotopes is approximately equal to one, but rises to three with increasing atomic number. The unstable isotopes of each element that lie to the left of the stable isotopes are *deficient* in neutrons, whereas those on the right side of the band of stable isotopes have an *excess* of neutrons. The existence of an excess or deficiency of neutrons in the nuclei of unstable atoms is reflected in the way they decay.

The phenomenon known to us as *radioactivity* consists of spontaneous nuclear transformations

that change the number of protons and neutrons in the nucleus. The process continues until a stable daughter is produced whose values of Z and N place it within the band of stable nuclides. The transformations are accompanied by the emission of alpha and beta particles and of gamma rays from the nucleus. Some isotopes of the light elements in Figure 16.1 actually disintegrate into alpha particles or emit protons or neutrons. Most radionuclides undergo *beta decay*, a much smaller number are subject to *alpha decay*, and a few of the heaviest isotopes decay by *spontaneous nuclear fission*.

a. Beta Decay

Beta decay can take place in three different ways, depending on whether the radionuclide has an excess or deficiency of neutrons. Radionuclides having an excess of neutrons emit a negatively charged beta particle from the nucleus. This has the effect of reducing the number of neutrons and increasing the number of protons:



where β^- is a negatively charged beta particle, which is identical to an extranuclear electron, and $\bar{\nu}$ is an antineutrino. The decay energy takes the form of kinetic energies of the beta particle and the antineutrino. The kinetic energy of the beta particle is complementary to that of the associated antineutrino, such that their sum is equal to the total decay energy. The product nuclide is the isotope of a different element because its atomic number differs from that of the parent. The nucleus of the product nuclide may have excess energy, which is emitted in the form of a gamma ray having a discrete energy. In some cases, the product nuclide may have two or more different excited states, which enables it to emit several gamma rays. If the product nuclide is unstable, it decays to form a second daughter by emitting another β^- particle, and so on, until at last a stable nuclide is formed. Each of the radionuclides in such a decay series emits β^- particles and antineutrinos with complementary kinetic energies equal to the total decay energy. In addition, each transition from one successive daughter to the next may be

masses into two gamma rays with a combined energy of 1.02 million electron volts (MeV). The neutrinos produced during positron decay have a range of kinetic energies complementary to those of the positrons, whereas those emitted during electron-capture decay are monoenergetic. Both escape into the universe and carry their energy with them.

A few radionuclides can decay by all three kinds of beta decay and, therefore, give rise to two different daughter products. For example, $^{40}_{19}\text{K}$ decays by positron decay and electron capture to stable $^{40}_{18}\text{Ar}$ and also decays by β^- (negatron) decay to stable $^{40}_{20}\text{Ca}$. Such *branched decay* also occurs among the unstable daughters of U and Th, some of which can decay both by β^- emission and by alpha decay.

b. Alpha Decay

Since the alpha particle is composed of two protons and two neutrons, it is identical to the nucleus of ^4_2He . The emission of an alpha particle lowers both the proton number and the neutron number of the product nucleus by two:



where P is the parent nuclide, D is the daughter, and Q is the total decay energy. Alpha decay may be followed by emission of gamma rays, which reduces the product nucleus to the ground state. Some of the kinetic energy of the alpha particles is imparted to the product nucleus as recoil energy, which may displace the nucleus from its position in the crystal lattice. The alpha particles are slowed by collisions and their energies are thereby converted into heat. These collisions also cause radiation damage in crystals containing alpha emitters.

Alpha decay is available to a large group of radionuclides having atomic numbers >58 and to a few nuclides of low atomic number, including ^5_2He , ^5_3Li , ^6_4Be , and others shown in Figure 16.1. The naturally occurring radioactive isotopes of U ($^{235}_{92}\text{U}$ and $^{238}_{92}\text{U}$) and Th ($^{232}_{90}\text{Th}$) decay to stable isotopes of Pb through series of short-lived radioactive daughters that form chains linked by alpha and beta decays. The decay chain of $^{238}_{92}\text{U}$ is illus-

trated in Figure 16.2. It contains 18 intermediate nuclides, all of which ultimately form stable $^{206}_{82}\text{Pb}$, regardless of the sequence of decay events. The isotopes $^{235}_{92}\text{U}$ and $^{232}_{90}\text{Th}$ each give rise to decay chains ending with $^{207}_{82}\text{Pb}$ and $^{208}_{82}\text{Pb}$, respectively.

All of the intermediate daughters of U and Th are short-lived radioactive nuclides that occur in the solar system only because they are continuously produced by decay of their long-lived parents. Some chemical elements including Po, At, Rn, Ra, and Pa occur only because their isotopes are members of these decay chains. The occurrence of Rn can be a hazard when it is inhaled since the alpha particles it emits can damage lung tissue.

c. Nuclear Fission

Some of the isotopes of U and of the transuranium elements decay by spontaneous fission. In this process the nucleus breaks into two fragments of unequal weight, which form the nuclei of two fission-product nuclides. These nuclides have an excess of neutrons, and therefore decay by emitting β^- particles and gamma rays until a stable nuclide is produced. Spontaneous fission liberates large amounts of energy and results in the emission of neutrons, alpha particles, and other fragments of nuclear material. The large amount of energy liberated by the spontaneous fission of $^{238}_{92}\text{U}$ in U-bearing minerals causes damage in the form of tracks about $10\text{ }\mu\text{m}$ in length. The tracks can be made visible by etching, so that they can be counted with a microscope, and are used to determine the ages and thermal histories of U-bearing minerals. The resulting *fission-track method* of dating has been described by Faure (1986).

Nuclear fission can also be *induced* when the nuclei of certain isotopes of U and Pu (plutonium) absorb neutrons or other nuclear particles. Induced fission of $^{235}_{92}\text{U}$ results in the release of neutrons, which can cause additional nuclei of $^{235}_{92}\text{U}$ to fission in a self-sustaining *chain reaction* that can liberate very large amounts of energy. In nuclear fission *reactors* such a chain reaction is maintained under controlled conditions and the heat generated by the reaction is used to generate electricity by means of conventional generators.

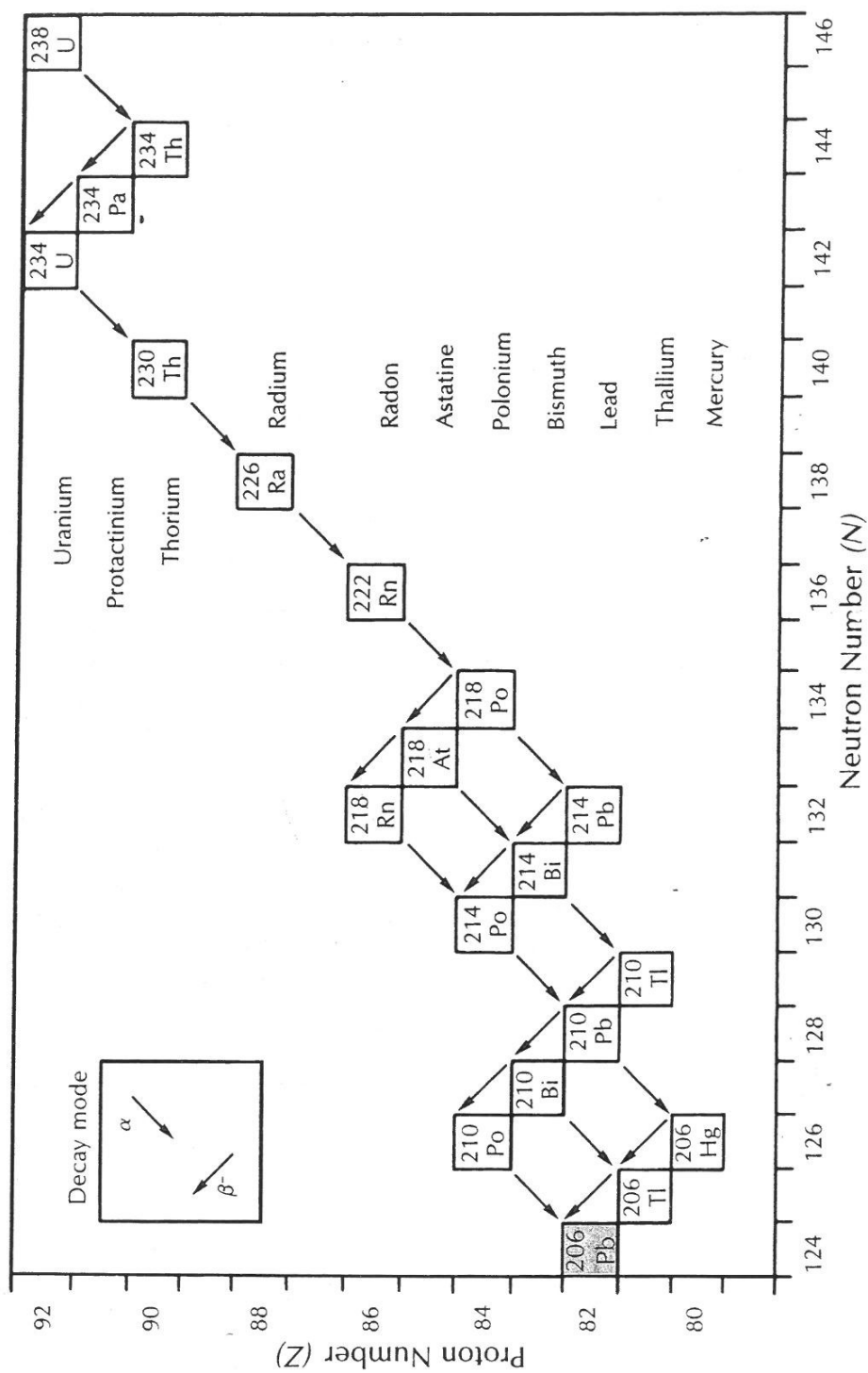


Figure 16.2 Chain of radioactive nuclides produced by decay of ^{238}U . The chain involves both alpha and beta (β^-) decay and branches repeatedly. However, all possible paths lead to stable ^{206}Pb . Note that this decay series includes 11 different chemical elements including Ra and Rn, which are important in environmental geochemistry.

The fission products accumulate in the *fuel rods*, which must be replaced when most of the $^{235}_{92}\text{U}$ has been consumed. The spent fuel rods are extremely radioactive because of the emission of β^- particles and gamma rays by the fission-product nuclides and their radioactive daughters. In addition, the spent fuel rods contain isotopes of Np (neptunium) and Pu (plutonium), that form from $^{238}_{92}\text{U}$ by neutron capture and subsequent beta decay leading to the formation of $^{239}_{94}\text{Pu}$ and $^{240}_{94}\text{Pu}$. These isotopes are *fissionable* and are used in the manufacture of nuclear weapons. Plutonium is also poisonous, and its dispersal creates an environmental health hazard (Chapter 24).

16.2 Law of Radioactivity

The *law of radioactivity* was discovered by Rutherford and Soddy (1902) and is expressed by the statement that the rate of decay of a radioactive nuclide is proportional to the number of atoms of that nuclide remaining at any time. If N is the *number* of atoms remaining, then:

$$-\frac{dN}{dt} = \lambda N \quad (16.5)$$

where λ is a proportionality constant known as the *decay constant* and the minus sign indicates that the rate of decay decreases with time. By integrating, we obtain:

$$\ln N = -\lambda t + C \quad (16.6)$$

where $\ln$ is the logarithm to the base e and C is the constant of integration. When $t = 0$, $C = \ln N_0$, where N_0 is the number of radioactive atoms at $t = 0$. Therefore, the integrated form of the equation is:

$$\ln N = -\lambda t + \ln N_0 \quad (16.7)$$

$$\ln \left(\frac{N}{N_0} \right) = -\lambda t \quad (16.8)$$

$$N = N_0 e^{-\lambda t} \quad (16.9)$$

It is convenient to define the *half-life* ($T_{1/2}$) as the time required for one half of a given number of radioactive atoms to decay. Therefore, if $t = T_{1/2}$,

then $N = N_0/2$. Substituting into equation 16.9, we obtain:

$$\frac{N_0}{2} = N_0 e^{-\lambda T_{1/2}} \quad (16.10)$$

$$\ln 1 - \ln 2 = -\lambda T_{1/2} \quad (16.11)$$

$$\ln 2 = \lambda T_{1/2} \quad (16.12)$$

$$T_{1/2} = \frac{\ln 2}{\lambda} \quad (16.13)$$

Every radionuclide has a characteristic decay constant that must be determined experimentally. This can be done by measuring its rate of decay at intervals over a period of time and by plotting the resulting *decay curve*. According to equation 16.5, the rate of disintegration (A) is equal to λN . Therefore, we can restate the law of radioactivity in terms of *disintegration rates* by multiplying both sides of equation 16.9 by λ :

$$\lambda N = \lambda N_0 e^{-\lambda t} \quad (16.14)$$

Since $A = \lambda N$, we obtain:

$$A = A_0 e^{-\lambda t} \quad (16.15)$$

Taking logarithms to the base e gives:

$$\ln A = \ln A_0 - \lambda t \quad (16.16)$$

Equation 16.16 is the equation of a straight line in the slope-intercept form where $\ln A_0$ is the intercept on the ordinate and $-\lambda$ is the slope. Evidently, the decay constant of a radionuclide can be determined from the slope of the straight line fitted to measurements of the disintegration rate at known intervals of time. The decay curve can also be plotted in terms of the number of radionuclides remaining (N) versus time measured in multiples of the half-life. Both forms of the decay curve are illustrated in Figure 16.3.

If the decay of a radionuclide gives rise to a stable daughter nuclide, we can say that the number of radiogenic daughter nuclides (D^*) that have accumulated at any time is equal to the number of parent nuclides that have decayed. Therefore,

$$D^* = N_0 - N \quad (16.17)$$

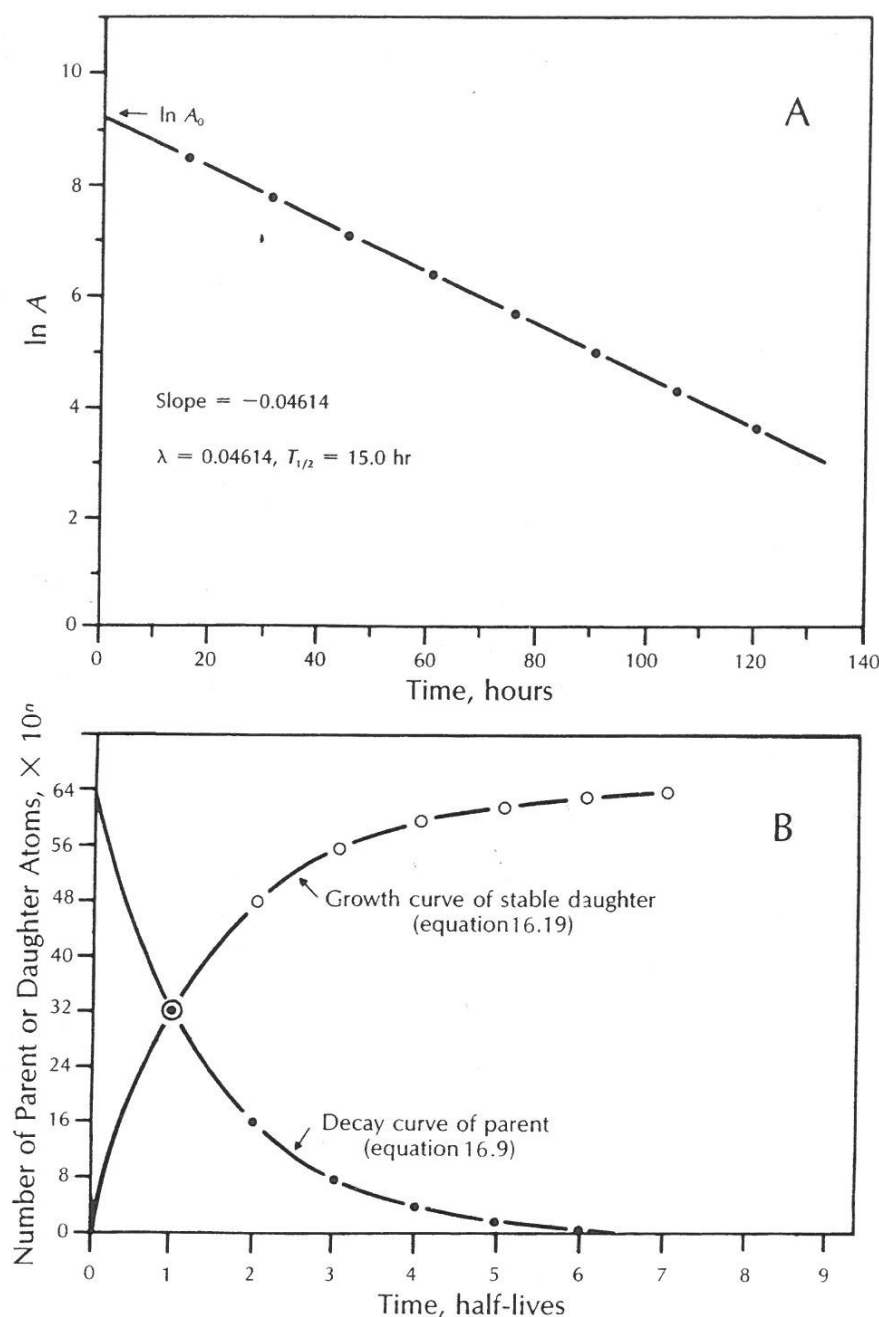


Figure 16.3 A. Decay of a radionuclide ($^{24}_{11}\text{Na}$) in ln-normal coordinates according to equation 16.16 where A is the disintegration rate and time t is measured in hours. The numerical value of the slope is equal to the decay constant λ from which the half-life $T_{1/2}$ can be calculated by means of equation 16.13.

B. Decay curve of a radionuclide and growth curve of its stable daughter in linear coordinates, where time is expressed in multiples of the half-life. Both curves can be constructed by using the definition of the half-life. Note that, in general, the number of parent atoms is reduced by a factor of 2^{-n} after n half-lives have elapsed.

where N_0 is the number of parent atoms at $t = 0$ and N is the number of parent atoms remaining at any time t . Substituting equation 16.9 yields:

$$D^* = N_0 - N_0 e^{-\lambda t} \quad (16.18)$$

$$D^* = N_0(1 - e^{-\lambda t}) \quad (16.19)$$

Equation 16.19 is the growth curve in Figure 16.3B, which relates the number of radiogenic daughters to the time that has elapsed and to the number of

parent atoms that were present initially when t was equal to zero. Notice that the decay curve of the parent and the growth curve of the stable daughter are complementary. This condition arises from equation 16.17, which is based on the assumption that the numbers of parent and daughter atoms change *only* as a result of decay and that no daughter or parent atoms are gained or lost by any other process.

The relationship between the radiogenic daughter and remaining parent atoms can also be expressed by replacing N_0 in equation 16.17 by:

$$N_0 = Ne^{\lambda t} \quad (16.20)$$

which we obtain from equation 16.9. Hence:

$$D^* = Ne^{\lambda t} - N \quad (16.21)$$

$$D^* = N(e^{\lambda t} - 1) \quad (16.22)$$

This is a more useful relationship than equation 16.19 because it relates the number of radiogenic daughter atoms (D^*) to the number of parent atoms (N) remaining, where both are *measurable* quantities. Therefore, equation 16.22 is the *geochronometry equation* used to determine the age of a mineral or rock that contains a long-lived naturally occurring radionuclide. The total number of radiogenic daughters (D) per unit weight of rock or mineral is:

$$D = D_0 + D^* \quad (16.23)$$

where D_0 is the number of daughters that entered a unit weight of mineral or rock at the time of its formation and D^* is the number of radiogenic daughters that formed by decay of the parent in the same unit weight of mineral or rock. A complete description of the number of daughter atoms is obtained by combining equations 16.22 and 16.23:

$$D = D_0 + N(e^{\lambda t} - 1) \quad (16.24)$$

Notice again that D and N are measurable quantities expressed in terms of numbers of atoms per unit weight of mineral or rock, D_0 is the number of daughters present *initially*, and t is the time elapsed since a sample of mineral or rock came

into existence. In other words, t is the age of the mineral or rock provided D and N changed only because of decay of the parent to the radiogenic daughter D^* .

In order to date a rock or mineral by means of equation 16.24, the measured concentrations of the parent and daughter elements must be recalculated into numbers of atoms of the appropriate isotopes by converting weights to numbers of moles and by using Avogadro's number and the abundances of the appropriate isotopes. These matters were discussed in Sections 6.3 and 6.4. However, in order to solve equation 16.24 for t we must also know the decay constant λ and the value of D_0 . The decay constants of the radionuclides have been measured and accurate values are available for most of them. However, how do we determine the value of D_0 ?

The number of daughter atoms present initially is determined in one of two ways. The least desirable but simplest method is to *assume* a value whose magnitude depends on the particular decay scheme and on the circumstances. For example, in the K-Ar method of dating we may assume that $D_0 = 0$ because Ar is a noble gas that is not likely to enter K-bearing minerals crystallizing from a magma. In the Rb-Sr method we may select a value for D_0 based on the assumption that the Sr originated from the mantle of the Earth and had an isotopic composition similar to that of Sr in midocean ridge basalt (MORB). The error caused by this assumption depends on the magnitude of the term $N(e^{\lambda t} - 1)$ in equation 16.24. If that term is much larger than D_0 , then the calculated date is not sensitive to the choice of D_0 . Therefore, as a rule, *old Rb-rich* minerals such as muscovite or biotite can be reliably dated using an assumed value of D_0 .

The preferred procedure is based on the analysis of several samples from the same body of rock such that all samples have the same age (t) and the same initial daughter content (D_0). For such a suite of samples D_0 and t are constants and equation 16.24 is a straight line in coordinates of D and N . This straight line is called an *isochron* because it is the locus of all points representing

mineral or rock samples that have the same age (t) and the same initial abundance of the daughter nuclide (D_0). The slope and intercept of the isochron are determined by the least-squares method using the analytical data from at least three *cogenetic* samples. The slope (m) of the isochron is related to the common age of the suite of cogenetic samples by:

$$m = e^{\lambda t} - 1 \quad (16.25)$$

and the age t is:

$$t = \frac{1}{\lambda} \ln(m + 1) \quad (16.26)$$

The *date* so calculated is the *age* of the rocks, provided the assumptions implied by the calculation are actually satisfied:

1. The number of parent and daughter atoms per unit weight of mineral or rock changed only by decay of the parent to the daughter.
2. The isotopic composition of the parent element was not altered by fractionation of isotopes at the time of formation of the mineral.

3. The decay constant of the parent is known accurately.
4. The isochron is not a mixing line.
5. The analytical data are accurate.

The units of time used in geochronology are years (a), thousands of years (ka), millions of years (Ma), or billions of years (Ga). In the ideal case described by these assumptions, the isochron technique yields not only the age of the samples but also indicates the abundance of the daughter isotope at the time of their formation (D_0). This value is of great interest because it contains information about the previous history of the daughter element before it was incorporated into the rocks in which it now resides (Dickin, 1995).

16.3 Methods of Dating

The Earth contains several naturally occurring long-lived radioactive isotopes, which are being used for dating by the methods outlined above. The parents and their daughters are listed in Table 16.1,

Table 16.1 Physical Constants Required for Dating Based on Naturally Occurring Long-Lived Radioactive Nuclides

Parent	Daughter	$T_{1/2},^a$ a	$\lambda,^a$ a^{-1}	Isotopic abundance of parent, %
$^{40}_{19}\text{K}$	$^{40}_{18}\text{Ar}$	1.19×10^{10}	0.581×10^{-10}	0.01167
$^{40}_{19}\text{K}$	$^{40}_{20}\text{Ca}$	1.40×10^{10}	4.962×10^{-10}	0.01167
$^{40}_{19}\text{K}$	$^{40}\text{Ar} + ^{40}\text{Ca}$	1.250×10^9	5.543×10^{-10}	0.01167
$^{87}_{37}\text{Rb}$	$^{87}_{38}\text{Sr}$	48.8×10^9	1.42×10^{-11}	27.8346
$^{147}_{62}\text{Sm}$	$^{143}_{60}\text{Nd}$	1.06×10^{11}	0.654×10^{-12}	15.0
$^{138}_{57}\text{La}$	$^{138}_{58}\text{Ce}$	2.70×10^{11}	2.57×10^{-12}	0.09
$^{138}_{57}\text{La}$	$^{138}_{56}\text{Ba}$	1.51×10^{11}	4.59×10^{-12}	0.09
$^{138}_{57}\text{La}$	$^{138}_{58}\text{Ce} + ^{138}_{56}\text{Ba}$	96.8×10^9	7.16×10^{-12}	0.09
$^{176}_{71}\text{Lu}$	$^{176}_{72}\text{Hf}$	35.7×10^9	1.94×10^{-11}	2.600
$^{187}_{75}\text{Re}$	$^{187}_{76}\text{Os}$	45.6×10^9	1.52×10^{-11}	62.602
$^{238}_{92}\text{U}$	$^{206}_{82}\text{Pb}$	4.468×10^9	1.55125×10^{-10}	99.2743
$^{235}_{92}\text{U}$	$^{207}_{82}\text{Pb}$	0.7038×10^9	9.8485×10^{-10}	0.7200
$^{232}_{90}\text{Th}$	$^{208}_{82}\text{Pb}$	14.010×10^9	4.9475×10^{-11}	100.00

^aa = year, derived from Latin word *annum*.

which also contains information about their decay constants, half-lives, and relevant isotopic abundances (Steiger and Jäger, 1977). Each method is applicable to those minerals and rocks that contain the parent element in significant concentrations and that retain the parent and daughter isotopes quantitatively. The applicability of the different dating methods is therefore governed, at least in part, by geochemical considerations.

a. The K–Ar Method

The positron and electron capture decay of radioactive $^{40}_{19}\text{K}$ to stable $^{40}_{18}\text{Ar}$ in K-bearing minerals causes the abundance of the radiogenic daughter to increase as a function of time in accordance with equation 16.24:

$$^{40}\text{Ar} = ^{40}\text{Ar}_0 + \left(\frac{\lambda_e}{\lambda}\right) ^{40}\text{K}(e^{\lambda t} - 1) \quad (16.27)$$

where ^{40}Ar and ^{40}K are the numbers of atoms of these isotopes per unit sample weight, λ_e is the decay constant for electron capture and positron decay ($\lambda_e = 0.581 \times 10^{-10} \text{ a}^{-1}$), and λ is the total decay constant of ^{40}K ($\lambda = 5.543 \times 10^{-10} \text{ a}^{-1}$). Note that $^{40}\text{Ar}_0 = 0$ and that the factor λ_e/λ is needed to identify the fraction of ^{40}K atoms that decay to ^{40}Ar .

The decay of $^{40}_{19}\text{K}$ to $^{40}_{20}\text{Ca}$ by β^- decay having a decay constant λ_β is expressed by a complementary equation:

$$^{40}\text{Ca} = ^{40}\text{Ca}_0 + \left(\frac{\lambda_\beta}{\lambda}\right) ^{40}\text{K}(e^{\lambda t} - 1) \quad (16.28)$$

This method of dating is rarely used because ^{40}Ca is the most abundant isotope of Ca (96.94%) and because its isotopic composition is difficult to measure by mass spectrometry.

The K–Ar method is widely used to date K-bearing minerals that retain ^{40}Ar quantitatively such as muscovite, biotite, and hornblende in igneous and metamorphic rocks. The dates calculated from equation 16.27 represent the time elapsed since the mineral being dated cooled through its *blocking temperature* below which ^{40}Ar is quantitatively retained. Therefore, K–Ar dates of slowly cooled minerals *underestimate* their ages. The K–Ar method has also been used

to date authigenic glauconite and illite in sedimentary rocks (Section 13.6) as well as whole-rock samples of glassy basalt. However, such applications require very careful selection of samples and are certainly not routine.

The $^{40}\text{Ar}/^{39}\text{Ar}$ method is a variation of the K–Ar method in which stable $^{39}_{19}\text{K}$ is partially converted to unstable but long-lived $^{39}_{40}\text{Ar}$ ($T_{1/2} = 269 \text{ a}$) by irradiating the sample with neutrons in a nuclear reactor. The sample is then heated stepwise and the $^{40}\text{Ar}/^{39}\text{Ar}$ ratio of Ar, released at each step, is used to calculate a date. The resulting spectrum of dates provides information about the cooling age of the sample as well as about the distribution of K and Ar within the mineral grains (Dalrymple and Lanphere, 1969).

b. The Rb–Sr Method

The β^- decay of naturally occurring $^{87}_{37}\text{Rb}$ to stable $^{87}_{38}\text{Sr}$ is the basis for the Rb–Sr method of dating. The geochronometry equation is written in terms of the isotopic ratio $^{87}\text{Sr}/^{86}\text{Sr}$, which is conveniently measured by mass spectrometry:

$$\frac{^{87}\text{Sr}}{^{86}\text{Sr}} = \left(\frac{^{87}\text{Sr}}{^{86}\text{Sr}}\right)_0 + \frac{^{87}\text{Rb}}{^{86}\text{Sr}}(e^{\lambda t} - 1) \quad (16.29)$$

The Rb–Sr method is used to date Rb-rich minerals such as muscovite, biotite, and K-feldspar in igneous and metamorphic rocks based on assumed values of the initial $^{87}\text{Sr}/^{86}\text{Sr}$ ratio, that is, 0.704 ± 0.001 , which is representative of Sr in the source regions of oceanic basalt. The resulting dates may indicate the time elapsed since the minerals cooled through their respective blocking temperatures for radiogenic ^{87}Sr , which lie *above* those for radiogenic ^{40}Ar . Therefore, Rb–Sr dates of slowly cooled minerals tend to be *older* than their K–Ar dates and are better estimates of their crystallization ages.

The Rb–Sr method is most often used to date suites of cogenetic igneous and high-grade metamorphic rocks by means of the *isochron* technique illustrated in Figure 16.4. Whole-rock samples of igneous and metamorphic rocks may retain Rb and Sr quantitatively even though radiogenic ^{87}Sr may have been redistributed among their constituent minerals by diffusion during

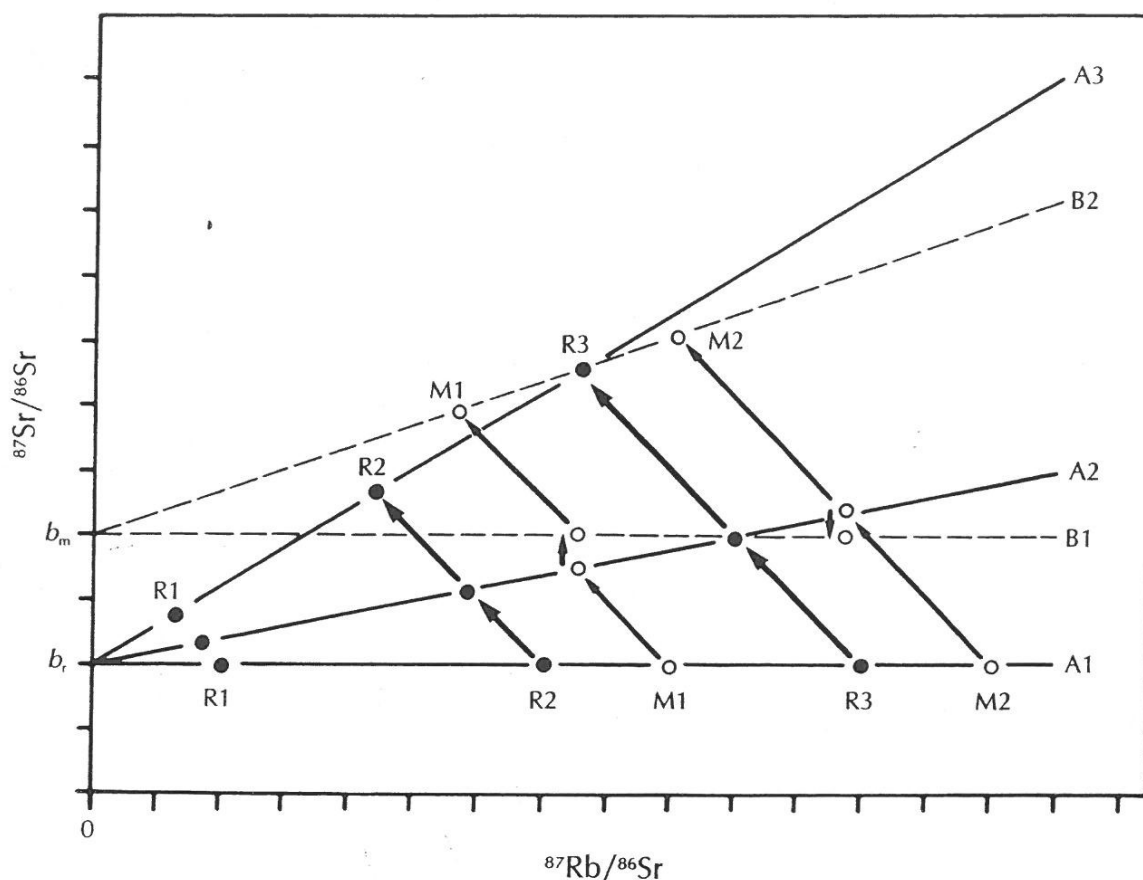


Figure 16.4 Rb–Sr isochron diagram for three whole-rock samples of igneous rocks (R1, R2, and R3) and two minerals (M1 and M2) of rock R3. The two axes in this diagram have equal scale to enhance the clarity of this illustration. Initially, all rock and mineral samples lie on isochron A1, which has a slope of zero and an intercept equal to b_r . As time passes, all rocks and minerals on A1 move along straight lines indicated by arrows and, at some time after crystallization, they form isochron A2. The slope of A2 is a measure of the amount of time that has passed since crystallization, but the intercept is still at b_r . The isotopic compositions of Sr in minerals M1 and M2 are changed at this time by diffusion of radiogenic ^{87}Sr caused by thermal metamorphism so that both minerals and R3 have the same $^{87}\text{Sr}/^{86}\text{Sr}$ ratio equal to b_m at the end of this episode. As a result, the rock R3 and its constituent minerals M1 and M2 now lie on isochron B1, but rock R3 also lies on isochron A2. After the passage of additional time, the whole-rock samples form isochron A3 whose slope is a function of the time elapsed since crystallization and whose intercept is still equal to b_r . However, minerals M1 and M2 now form isochron B2 whose slope is less than that of isochron A3 and measures the time elapsed since their isotopic reequilibration. The intercept of isochron B2 at b_m is greater than b_r because it is the $^{87}\text{Sr}/^{86}\text{Sr}$ ratio of R3 and its constituent minerals after isotopic reequilibration. This hypothetical example suggests that whole-rock Rb–Sr isochrons can yield the crystallization age of igneous and high-grade metamorphic rocks, whereas their constituent minerals form isochrons that indicate the time elapsed since the last episode of isotopic reequilibration.

thermal metamorphism. The date derived from the slope of a whole-rock isochron is therefore regarded as a reliable indicator of the crystallization age of igneous and metamorphic rocks. The intercept is the initial $^{87}\text{Sr}/^{86}\text{Sr}$ ratio, which reflects the geochemical history of the Sr before it was incorporated into the rocks that form the isochron. The initial $^{87}\text{Sr}/^{86}\text{Sr}$ is, therefore, an important parameter for the study of the petrogenesis of igneous rocks and of the geochemical differentiation of the continental crust from the underlying mantle (Faure and Powell, 1972).

c. The Sm–Nd, La–Ce, and Lu–Hf Methods

These rare earth geochronometers are useful for dating Ca-rich igneous and high-grade metamorphic rocks of Precambrian age. The relevant equations (equations 16.30–16.32), follow directly from equation 16.24 and are written in terms of isotopic ratios:

$$\frac{^{143}\text{Nd}}{^{144}\text{Nd}} = \left(\frac{^{143}\text{Nd}}{^{144}\text{Nd}} \right)_0 + \frac{^{147}\text{Sm}}{^{144}\text{Nd}} (e^{\lambda t} - 1) \quad (16.30)$$

$$\frac{^{138}\text{Ce}}{^{140}\text{Ce}} = \left(\frac{^{138}\text{Ce}}{^{140}\text{Ce}} \right)_0 + \left(\frac{\lambda_\beta}{\lambda} \right) \frac{^{138}\text{La}}{^{140}\text{Ce}} (e^{\lambda t} - 1) \quad (16.31)$$

$$\frac{^{176}\text{Hf}}{^{177}\text{Hf}} = \left(\frac{^{176}\text{Hf}}{^{177}\text{Hf}} \right)_0 + \frac{^{176}\text{Lu}}{^{177}\text{Hf}} (e^{\lambda t} - 1) \quad (16.32)$$

All three methods are used to construct isochrons by analysis of suites of cogenetic mafic or even ultramafic rocks. These geochronometers can be used to date rocks that are not datable by the K–Ar and Rb–Sr methods. In addition, the initial isotope ratios are important for studies of the petrogenesis of igneous rocks and of crustal evolution.

d. The Re–Os Method

The Re–Os method has become easier to use because of improvements in the isotopic analysis of Os (Creaser et al., 1991). The average concentrations of Re and Os in rocks are very low and amount to only fractions of parts per billion. Both elements are siderophile and are strongly concentrated in the Fe–Ni phases of meteorites. In addition, Re occurs in *molybdenite* (MoS_2), whereas

Os is concentrated with other Pt-group metals in the naturally occurring alloy *osmiridium*.

The geochronometry equation takes the form:

$$\frac{^{187}\text{Os}}{^{186}\text{Os}} = \left(\frac{^{187}\text{Os}}{^{186}\text{Os}} \right)_0 + \frac{^{187}\text{Re}}{^{186}\text{Os}} (e^{\lambda t} - 1) \quad (16.33)$$

It is used to construct isochrons for dating of meteorites, ultramafic rocks, and molybdenite-bearing ore samples. An advantage of the Re–Os decay scheme is that crustal rocks have high Re/Os ratios. For example, the average Re/Os ratio of tholeiite basalts is 28 and that of granites is about 10. This means that crustal rocks of Precambrian age have high $^{187}\text{Os}/^{186}\text{Os}$ ratios, which is useful to igneous petrologists who study the contamination of basalt magma by assimilation of crustal rocks.

e. The U, Th–Pb Methods

The decay of ^{238}U , ^{235}U , and ^{232}Th runs through separate series of intermediate radioactive daughters to stable isotopes of Pb. These decay chains can achieve a condition of secular equilibrium when the rates of decay of the parents are imposed on all of their intermediate daughters. Under these conditions the geochronometry equations are:

$$\frac{^{206}\text{Pb}}{^{204}\text{Pb}} = \left(\frac{^{206}\text{Pb}}{^{204}\text{Pb}} \right)_0 + \frac{^{238}\text{U}}{^{204}\text{Pb}} (e^{\lambda t} - 1) \quad (16.34)$$

$$\frac{^{207}\text{Pb}}{^{204}\text{Pb}} = \left(\frac{^{207}\text{Pb}}{^{204}\text{Pb}} \right)_0 + \frac{^{235}\text{U}}{^{204}\text{Pb}} (e^{\lambda t} - 1) \quad (16.35)$$

$$\frac{^{208}\text{Pb}}{^{204}\text{Pb}} = \left(\frac{^{208}\text{Pb}}{^{204}\text{Pb}} \right)_0 + \frac{^{232}\text{Th}}{^{204}\text{Pb}} (e^{\lambda t} - 1) \quad (16.36)$$

The three decay schemes can yield three independent age determinations of minerals or rocks containing both U and Th, provided they did not gain or lose U, Th, or Pb after the date of their formation. In most cases, however, the dates indicated by these geochronometers do *not* agree with each other and are said to be *discordant*. Most commonly, the discordance results from loss of Pb or of intermediate daughters. The loss of daughter nuclides from U, Th-bearing minerals is caused in part by the radiation damage resulting from the energy released during alpha decay.

In such cases, a fourth date can be calculated by combining equations 16.34 and 16.35 for the U–Pb geochronometers:

$$\frac{\frac{^{207}\text{Pb}}{^{204}\text{Pb}} - \left(\frac{^{207}\text{Pb}}{^{204}\text{Pb}}\right)_0}{\frac{^{206}\text{Pb}}{^{204}\text{Pb}} - \left(\frac{^{206}\text{Pb}}{^{204}\text{Pb}}\right)_0} = \frac{^{235}\text{U}}{^{238}\text{U}} \left(\frac{e^{\lambda_2 t} - 1}{e^{\lambda_1 t} - 1} \right) \quad (16.37)$$

This equation represents a family of straight lines in coordinates of $^{207}\text{Pb}/^{204}\text{Pb}$ (y) and $^{206}\text{Pb}/^{204}\text{Pb}$ (x) that intersect in a point whose coordinates are the initial $^{207}\text{Pb}/^{204}\text{Pb}$ and $^{206}\text{Pb}/^{204}\text{Pb}$ ratios. Therefore, the left side of equation 16.37 is the slope of these straight lines, which are known as *Pb–Pb isochrons*:

$$\frac{\frac{^{207}\text{Pb}}{^{204}\text{Pb}} - \left(\frac{^{207}\text{Pb}}{^{204}\text{Pb}}\right)_0}{\frac{^{206}\text{Pb}}{^{204}\text{Pb}} - \left(\frac{^{206}\text{Pb}}{^{204}\text{Pb}}\right)_0} = \left(\frac{^{207}\text{Pb}}{^{206}\text{Pb}}\right)^* = \text{slope} \quad (16.38)$$

where the asterisk (*) indicates that the slope of the Pb–Pb isochrons is the ratio of the *radiogenic* Pb isotopes. Equation 16.37 relates the ratio of the radiogenic Pb isotopes to the age t of the rock or mineral based *only* on the isotopic composition of Pb. However, the equation is transcendental and cannot be solved for t by ordinary algebraic methods. Solutions can be obtained most conveniently by interpolating in a table of calculated $(^{207}\text{Pb}/^{206}\text{Pb})^*$ ratios and t (Faure, 1986, Table 18.3, p. 290). In cases where equations 16.34–16.36 yield discordant dates, the *207/206 date* derived from equation 16.37 is commonly older than the U–Pb and Th–Pb dates because the Pb–Pb date is not affected by recent Pb loss that does not change the isotopic composition of the remaining Pb.

Although the number of known U- and Th-bearing minerals is very large, only a few are suitable for dating by the U, Th–Pb methods because they retain radiogenic Pb and because they are *common*. Zircon (ZrSiO_4) meets these criteria better than all other minerals, although monazite (rare earth phosphate) and apatite (Ca phosphate) are used occasionally. Zircon permits the replacement of Zr^{4+} by U^{4+} and Th^{4+} in accordance with Goldschmidt's first rule (Table 7.3 and

Section 8.1) because their radii in sixfold coordinated sites are similar ($\text{Zr}^{4+} = 0.80 \text{ \AA}$, $\text{U}^{4+} = 0.97 \text{ \AA}$, and $\text{Th}^{4+} = 1.08 \text{ \AA}$). However, zircon strongly *excludes* Pb, which can form both Pb^{2+} ($1.26 \text{ \AA}$) and Pb^{4+} ($0.86 \text{ \AA}$) ions. Since Pb is excluded from zircon, it must occur as Pb^{2+} , which cannot replace Zr^{4+} because it is too large and its charge differs too much from that of Zr^{4+} . Therefore, Goldschmidt's rules help to explain why zircon crystals exclude Pb but accept U and Th, thereby making zircon one of the most useful geological chronometers.

However, not even zircon is perfectly retentive and, in most cases, U, Th–Pb dates of zircons are *discordant*. This problem was solved brilliantly by Wetherill (1956) by means of the *concordia diagram*, which is constructed by calculating the $^{206}\text{Pb}^*/^{238}\text{U}$ and $^{207}\text{Pb}^*/^{235}\text{U}$ ratios of U–Pb systems that yield concordant dates. From equation 16.34 we have:

$$\frac{\frac{^{206}\text{Pb}}{^{204}\text{Pb}} - \left(\frac{^{206}\text{Pb}}{^{204}\text{Pb}}\right)_0}{\frac{^{238}\text{U}}{^{204}\text{Pb}}} = \frac{^{206}\text{Pb}^*}{^{238}\text{U}} = e^{\lambda_1 t} - 1 \quad (16.39)$$

Similarly,

$$\frac{^{207}\text{Pb}^*}{^{235}\text{U}} = e^{\lambda_2 t} - 1 \quad (16.40)$$

Compatible sets of $^{206}\text{Pb}^*/^{238}\text{U}$ and $^{207}\text{Pb}^*/^{235}\text{U}$ ratios, calculated from equations 16.39 and 16.40 for selected values of t , are the coordinates of points that form the *concordia curve*. In other words, the concordia curve is the locus of all points representing U–Pb systems that yield concordant U–Pb dates. Any U-bearing rock or mineral that does *not* plot on concordia yields *discordant* U–Pb dates.

Zircon crystals or other U–Pb systems that remain closed to U and Pb, and to the intermediate daughters, move along the concordia curve in Figure 16.5 as they age. When they lose Pb as a result of an episode of thermal metamorphism, they leave concordia and move along a straight line toward the origin. When all accumulated radiogenic Pb is lost, the $^{206}\text{Pb}/^{238}\text{U}$ and

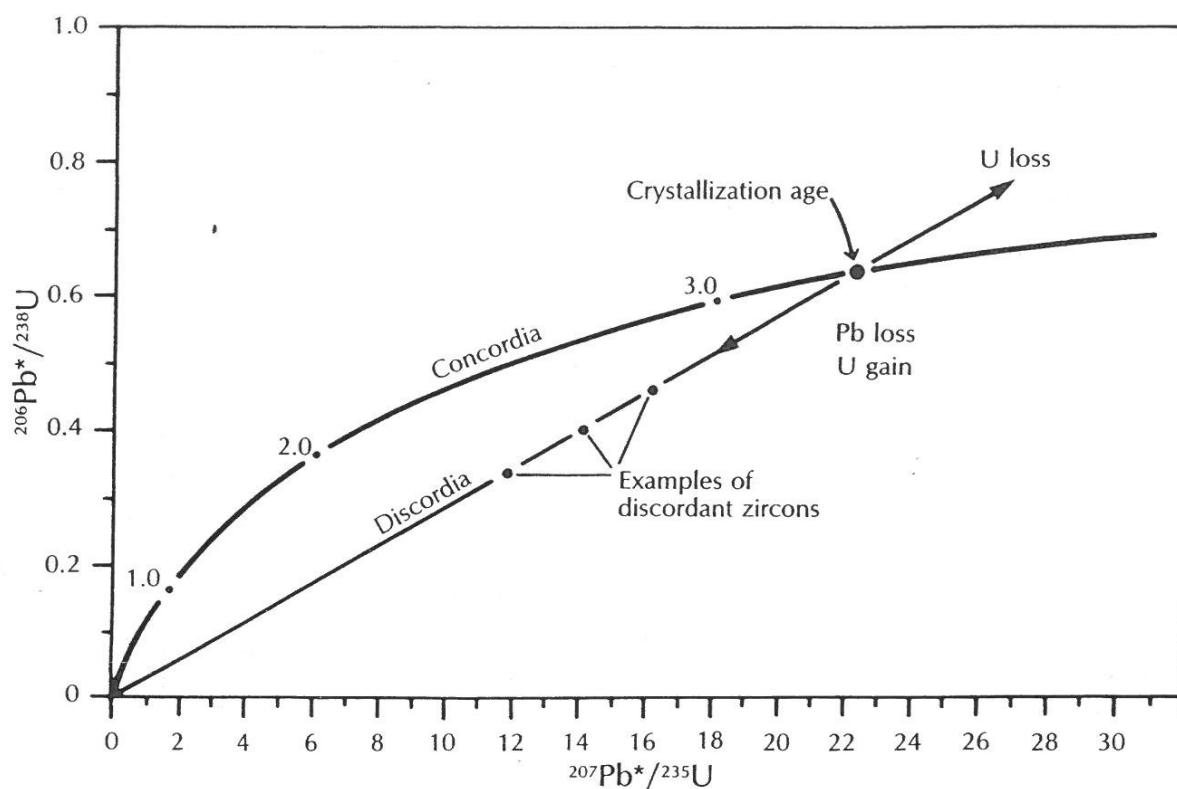


Figure 16.5 Concordia diagram used in dating minerals such as zircon, monazite, apatite, and others by the U–Pb method. The coordinates of points on concordia are generated by solving equations 16.38 and 16.39 for selected values of t . During episodic Pb loss or U gain, minerals are displaced from concordia and move along the discordia line toward the origin. Zircons, or other U-bearing minerals, which yield discordant U–Pb dates by solution of equations 16.33 and 16.34, are represented by points that define discordia. The upper point of intersection of discordia with concordia is the crystallization age of the zircons that define discordia. Loss of U causes U–Pb minerals to be represented by points on discordia *above* concordia. The effect of Pb gain is not predictable unless the isotopic composition of the added Pb is known. U–Th–Pb concordias are possible but have been rarely used because differences in the geochemical properties of U and Th may produce discordias that do not pass through the origin.

$^{207}\text{Pb}/^{235}\text{U}$ ratios approach zero along a chord directed toward the origin. When only a fraction of Pb is lost from zircons, they are represented by points on the straight line. Such U–Pb systems yield discordant dates, which is why the chord is called *discordia*.

When zircon crystals from a large sample of granitic rock are sorted according to their color, magnetic properties, or shape prior to analysis, they commonly form points that define a straight line on the concordia diagram. The upper point of intersection of this discordia line with concordia

yields the time elapsed since crystallization because, if the zircons had not lost Pb, they would have resided on concordia at that point. The lower point of intersection *may* be the time elapsed since the end of episodic Pb loss, because this is where zircon crystals would now be located on concordia if they had lost all of the accumulated Pb during the episode of thermal metamorphism (see Section 19.5).

The significance of the lower point of intersection of discordia with concordia is uncertain because Pb loss may be *continuous* rather than

episodic, or it may be caused by the relaxation of lithostatic pressure as a result of erosion of the overlying rocks following crustal uplift. In addition, zircon crystals and other U-bearing minerals may suffer not only loss of Pb but may also lose or gain U. Gain of U has the same effect on U-Pb systems as loss of Pb, whereas loss of U causes them to move *away* from the origin and places them on discordias *above* concordia. This kind of behavior is sometimes seen in monazite and other U-bearing phosphate minerals. In such cases, the age of the minerals is indicated by the upper point of intersection as before, but the lower point of intersection does not have a unique interpretation.

Use of highly selected and purified samples consisting of only a small number of zircon crystals (or fragments of a single crystal) can reduce the statistical error of ages determined by this method to ± 2 million years (Ma) for zircons whose age is 2700 Ma (Krogh, 1982a,b).

The decay of U and Th to stable isotopes of Pb permits still more elaborate interpretations that go beyond the limits of this book. We have already mentioned the fission-track method of dating, which arises from the spontaneous fission decay of ^{238}U . In addition, the growth of radiogenic Pb in the Earth is the subject of *plumbology*, which seeks to interpret the isotopic composition of Pb in Pb-bearing minerals that do not contain U or Th. Therefore, plumbology is concerned with the study of Pb in ore deposits containing galena (PbS) and minerals of other metals (Stacey and Kramers, 1975). Before we leave this subject, however, we consider briefly the so-called U-series disequilibrium methods of dating.

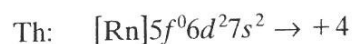
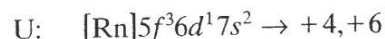
f. The Ionium (^{230}Th) Methods

The daughters of ^{238}U , shown in Figure 16.2, form a decay chain that can be broken by geological processes because the daughters of U are isotopes of different elements and therefore have different geochemical properties. Two different situations may arise:

1. A member of the decay chain is separated from its parent and therefore "dies" at a rate determined by its half-life.
2. A member of the decay chain is separated from its daughter and subsequently regenerates its own daughter.

Both conditions occur between ^{234}U and its daughter ^{230}Th (ionium) in the decay series of ^{238}U .

Uranium can assume valences of +4 and +6, whereas Th has only one valence of +4. The reason for this difference is that U atoms ($Z = 92$) have two more electrons than Th atoms ($Z = 90$). Their electron configurations (Table 5.2) are:



The tetravalent (+4) ions of U and Th form compounds that are insoluble in water. In addition, their high ionic potentials (charge-to-radius ratio) cause them to be strongly adsorbed on the charged surface of small particles. However, in the presence of molecular O at or near the surface of the Earth, U is oxidized to the hexavalent state (+6) and forms the *uranyl* ion (UO_2^{2+}), which has a smaller ionic potential than U^{4+} and therefore is not scavenged from aqueous solutions by small particles as is Th^{4+} . Therefore, when U-bearing minerals decompose as a result of chemical weathering, ^{234}U forms the uranyl ion and remains in solution, whereas ^{230}Th forms a tetravalent ion and is strongly adsorbed on sedimentary particles. As a result, ^{230}Th (ionium) is separated from its parent by being concentrated in sediment deposited in the oceans or in lakes.

After deposition of the sediment, the adsorbed "excess" ^{230}Th decays by emitting alpha particles with its characteristic half-life ($T_{1/2} = 7.52 \times 10^4 \text{ a}$, $\lambda = 9.217 \times 10^{-6} \text{ a}^{-1}$) in accordance with the law of radioactivity. The number of ^{230}Th atoms remaining in a unit weight of dry sediment is measured by observing the rate of disintegration (A) using an alpha spectrometer of known efficiency. When samples are taken from a piston core, the ^{230}Th activity (^{230}Th)_A of a particular sample is related to its age (t) by equation 16.15:

$$({}^{230}\text{Th})_A = ({}^{230}\text{Th})_A^0 e^{-\lambda t} \quad (16.41)$$

where $({}^{230}\text{Th})_A^0$ is the activity of ${}^{230}\text{Th}$ at the top of the core. For practical reasons, the activity of ${}^{230}\text{Th}$ is divided by the activity of adsorbed ${}^{232}\text{Th}$, whose activity is virtually independent of time because of its long half-life, (Table 16.1). When the rate of sedimentation s is defined as:

$$s = \frac{h}{t} \quad (16.42)$$

where h is the depth below the top of the core, $t = h/s$, and the ratio of the activities of the Th isotopes is:

$$\left(\frac{{}^{230}\text{Th}}{{}^{232}\text{Th}}\right)_A = \left(\frac{{}^{230}\text{Th}}{{}^{232}\text{Th}}\right)_A^0 e^{-\lambda h/s} \quad (16.43)$$

We linearize the equation by taking natural logarithms:

$$\ln \left(\frac{{}^{230}\text{Th}}{{}^{232}\text{Th}}\right)_A = \ln \left(\frac{{}^{230}\text{Th}}{{}^{232}\text{Th}}\right)_A^0 - \frac{\lambda h}{s} \quad (16.44)$$

Figure 16.6 illustrates how sediment samples taken at different depths from a core generate an array of points in coordinates of $\ln ({}^{230}\text{Th}/{}^{232}\text{Th})_A$ and h to which a straight line can be fitted. The intercept of that line with the $\ln ({}^{230}\text{Th}/{}^{232}\text{Th})_A$ (y) axis is the ${}^{230}\text{Th}/{}^{232}\text{Th}$ activity ratio at the sediment–water interface and the slope of the line is:

$$m = -\frac{\lambda}{s} \quad (16.45)$$

Since the slope also has a negative value, it follows that the sedimentation rate is:

$$s = \frac{\lambda}{m} \quad (16.46)$$

This method of dating works best when it is applied to sediment samples taken from known depths below the sediment–water interface. It can yield not only the ages of individual samples based on the extrapolated value of $({}^{230}\text{Th}/{}^{232}\text{Th})_A^0$ but also indicates sedimentation rates. The method is limited by the eventual regrowth of ${}^{230}\text{Th}$ from small amounts of ${}^{234}\text{U}$ adsorbed on the sediment particles, which causes data points to

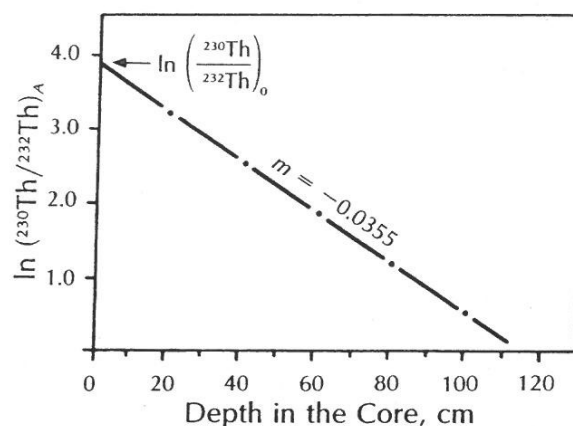


Figure 16.6 Decay of excess ${}^{230}\text{Th}$ in a core of detrital sediment at increasing depths below the water–sediment interface according to equation 16.44. The slope of the line fitted to the hypothetical data points is $m = -0.0355$. According to equation 16.46, the sedimentation rate is $s = \lambda/m = (9.217 \times 10^{-6})/0.0355 = 2.60 \times 10^{-4}$ cm/a or 2.60 mm/ 10^3 yr. This value is fairly typical of the rate of deposition of detrital sediment in the oceans.

deviate from the linear relationship required by equation 16.44.

The differences in geochemical properties of U and Th come into play in yet another way. The water-soluble uranyl ion (UO_2^{2+}) may enter calcite or aragonite by replacement of Ca^{2+} , whereas Th^{4+} is strongly excluded. Therefore, freshly deposited calcium carbonate contains ${}^{234}\text{U}$ that has been separated from its daughter ${}^{230}\text{Th}$. The daughter is then regenerated in the calcite or aragonite at a rate determined by the difference between the rate of decay of its parent and its own decay rate. The differential equation governing the decay of a radionuclide to a radioactive daughter was solved by Bateman (1910). When the solution of that equation is adapted to the decay of ${}^{230}\text{Th}$ that is forming by decay of ${}^{234}\text{U}$, we obtain (Faure, 1986, ch. 21, sec. 3):

$$({}^{230}\text{Th})_A = ({}^{238}\text{U})_A (1 - e^{-\lambda_{230}t}) \quad (16.47)$$

This equation is similar to equation 16.19, which relates the number of daughters to the number of

parent atoms present initially. In equation 16.47 the parent is ^{238}U , which is assumed to be in secular equilibrium with its daughter ^{234}U . Since ^{238}U has a very long half-life ($T_{1/2} = 4.468 \times 10^9$ a, Table 16.2), its activity is virtually constant on a time scale of 1 million years. Therefore, the growth curve of the activity of ^{230}Th in calcium carbonate in Figure 16.7 has the same shape as the one in Figure 16.3B, which is a plot of equation 16.19.

The ^{234}U - ^{230}Th method of dating is useful for dating calcium carbonate deposited in the oceans in the past 150,000 years or less. It has been used to date coral reefs related to sea-level fluctuations and the results have supported the *Milankovitch hypothesis* that the global climate is affected by periodic variations in the amount and timing of solar radiation received by the Earth

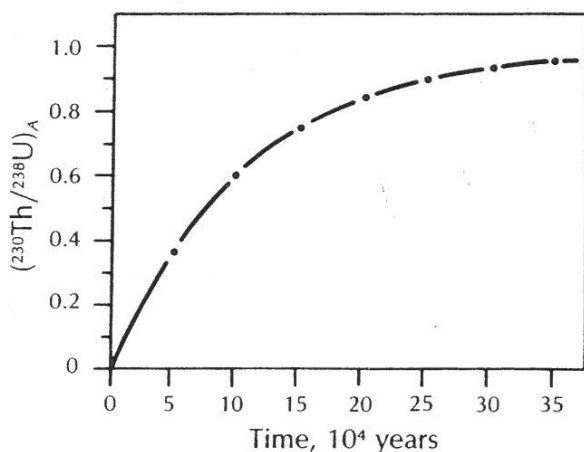


Figure 16.7 Growth of ^{230}Th by decay of ^{234}U , which is in secular radioactive equilibrium with ^{238}U , in accordance with equation 16.47. During the formation of calcium carbonate from aqueous solution U is admitted as the uranyl ion (UO_2^{2+}), whereas Th is excluded. Therefore, the decay chain of ^{238}U is thereby broken and ^{234}U in the carbonate regenerates the daughter ^{230}Th until their activity ratio reaches a value of one and secular equilibrium is reestablished. The diagram assumes that ^{234}U is in equilibrium with ^{238}U , but in reality the oceans and water on the continents contain excess ^{234}U because it is leached preferentially from U-bearing minerals. A derivation of the equation representing cases where $(^{234}\text{U}/^{238}\text{U})_A > 1$ was given by Faure (1986, pp. 370-371).

because of the dynamics of its orbit around the Sun (Broecker et al., 1968).

Equation 16.47 is inexact because ^{234}U is actually *not* in secular equilibrium with ^{238}U in water on the surface of the Earth. The reason is that ^{234}U is leached preferentially from U-bearing minerals because it resides in lattice sites that were damaged by the decay of ^{238}U and of the two daughters that precede ^{234}U (Figure 16.2). As a result, the disintegration rate of ^{234}U in the oceans is 15% *greater* than that of ^{238}U . The discrepancy is even greater in groundwater, where the activity ratio $^{234}\text{U}/^{238}\text{U}$ may be as high as 10. A complete description of the growth of ^{230}Th in systems containing excess ^{234}U can be found in Faure (1986, p. 370) and in the references cited therein.

The decay of ^{238}U and ^{235}U through chains of unstable daughters has permitted the development of several other geochronometers. All of them are used to study the rates of sedimentation and other geological processes on the surface of the Earth. These applications of *radiochemistry* to geology have grown into a large subdivision of isotope geochemistry.

16.4 Cosmogenic Radionuclides

A large group of radionuclides is produced by nuclear reactions between cosmic rays and the stable atoms of the atmosphere and in rocks exposed at the surface of the Earth. Cosmic rays originate primarily from the Sun and consist predominantly of highly energetic protons. The products of these nuclear reactions may be removed from the atmosphere by meteoric precipitation or they may remain in the atmosphere for long periods of time, depending on their chemical properties. Many of the product nuclides have half-lives measured in hours or days and therefore are not useful for geologic research. However, a surprisingly large number of these nuclides decay sufficiently slowly to be of interest in geology.

The radionuclides listed in Table 16.2 are becoming increasingly useful because of dramatic improvements in the equipment used to detect them. The traditional measurement technique is low-level counting using heavily shielded radia-

Table 16.2 Physical Constants of Some Cosmogenic Radionuclides

Nuclide	$T_{1/2}, \text{a}$	λ, a^{-1}
^3_1H (tritium)	12.26	5.653×10^{-2}
$^{10}_4\text{Be}$	1.5×10^6	0.462×10^{-6}
$^{14}_6\text{C}$	5730	0.1209×10^{-3}
$^{26}_{13}\text{Al}$	0.716×10^6	0.968×10^{-6}
$^{32}_{14}\text{Si}$	276	0.251×10^{-2}
$^{36}_{17}\text{Cl}$	0.308×10^6	2.25×10^{-6}
$^{39}_{18}\text{Ar}$	269	0.257×10^{-2}
$^{55}_{25}\text{Mn}$	3.7×10^6	0.187×10^{-6}
$^{59}_{28}\text{Ni}$	8×10^4	0.086×10^{-4}
$^{81}_{36}\text{Kr}$	0.213×10^6	3.25×10^{-6}

tion detectors. In recent years, ultrasensitive high-energy tandem mass spectrometers, which literally count ions, have greatly improved the sensitivity with which cosmogenic radionuclides can be detected.

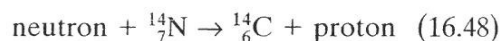
The atmospheric inventories of the cosmogenic radionuclides depend on the rates at which they are produced and on the rates at which they decay or are removed from the atmosphere by meteoric precipitation or by other processes. Once a particular radionuclide has been removed from its atmospheric reservoir, decay becomes the dominant process and its activity declines with time in accordance with equation 16.15. The measured disintegration rate of a particular cosmogenic radionuclide in a geological sample can be used to solve equation 16.15 for t , provided the initial disintegration rate is known or can be derived from a set of data.

a. The Carbon-14 Method

The carbon-14 method of dating was developed in the late 1940s by a group of scientists working at the University of Chicago. The group included W. F. Libby, J. R. Arnold, E. C. Anderson, and A. V. Grosse, who successfully demonstrated the presence of ^{14}C in methane from a sewage-treatment plant. The first age determinations of archaeological objects of known age were made by

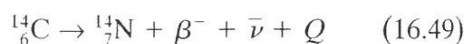
Arnold and Libby (1949). Two years later, Anderson and Libby (1951) reported that the specific disintegration rate of ^{14}C in modern plants is constant worldwide and has an average value of 15.3 ± 0.5 disintegrations per minute per gram (dpm/g) of carbon. This value was later revised to 13.56 ± 0.07 dpm/g of carbon. These measurements therefore established the feasibility of the ^{14}C method of dating, which Libby described in a book published in 1952. The method was widely acclaimed and earned W. F. Libby the Nobel Prize in chemistry in 1960. In 1959 a journal was founded called *Radiocarbon* in which all age determinations and many research papers dealing with the distribution of ^{14}C continue to be published (Mook, 1980).

The dominant nuclear reaction producing ^{14}C in the atmosphere involves the interaction of neutrons, derived from cosmic rays, with stable $^{14}_7\text{N}$:



The radioactive ^{14}C so produced is incorporated into CO_2 molecules and is rapidly mixed throughout the atmosphere. From there it is absorbed by green plants during photosynthesis and enters cellulose and other carbon compounds manufactured by them. It is transmitted further up the food chain by herbivorous and carnivorous animals so that virtually all living organisms contain ^{14}C at about the same specific concentration expressed in terms of disintegrations per minute per gram of carbon.

As long as a plant is alive, it continues to absorb ^{14}C from the atmosphere. However, when it dies, the concentration of ^{14}C in its tissue begins to decline at a rate dependent on the half-life of ^{14}C as shown in Figure 16.8. The decay of ^{14}C takes place by β^- emission to stable $^{14}_7\text{N}$:



where the symbols have the same meaning as in equation 16.1. The half-life of ^{14}C is 5730 years. However, ^{14}C dates are based on a value of 5568 years for the decay constant.

The basic principle and measurement technique of the carbon-14 method of dating are quite simple. The specific activity of ^{14}C in a sample containing biogenic carbon is measured by converting

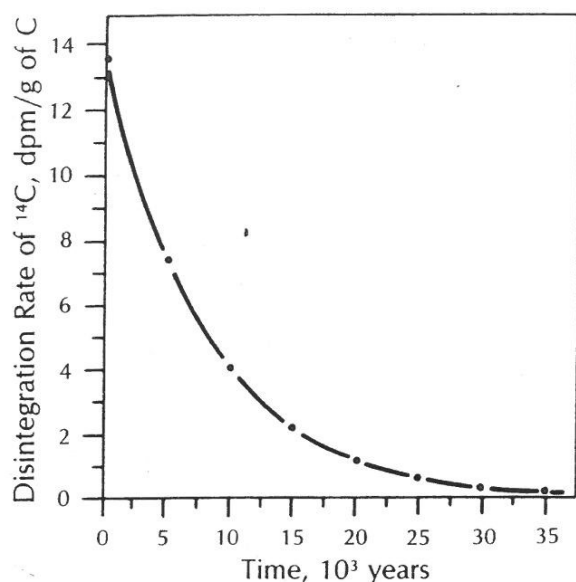


Figure 16.8 Decay curve of ^{14}C in accordance with equation 16.15. The initial decay rate of ^{14}C is 13.56 ± 0.07 dpm/g of C. The ^{14}C date is calculated from equation 16.50 based on the measured disintegration rate of the remaining ^{14}C . This method can be used to date plant remains and calcium carbonate formed in equilibrium with CO_2 of the atmosphere. The limit of the technique based on radiation detectors is about 35,000 years but can be doubled by using ultrasensitive mass spectrometers. Although the carbon-14 dating method is simple in principle, it is complex in reality because the initial activity of ^{14}C has varied with time in the past and because of fractionation of the isotopes of C.

the carbon to CO_2 , or another carbon compound, and by counting it in a well-shielded *ionization chamber* or *scintillation counter*. The measured activity (A) is substituted into equation 16.15, which is solved for t using the assumed initial activity of ^{14}C (A_0):

$$t = \frac{1}{\lambda} \ln \frac{A_0}{A} \quad (16.50)$$

The materials suitable for dating include charcoal and wood, seeds and nuts, grasses, cotton cloth, paper, leather, and peat. In addition, calcium carbonate in mollusk shells and in marine or lacustrine sediment can be dated because such deposits

contain CO_2 derived partly from the atmosphere. The initial activity of ^{14}C in calcium carbonate samples, however, is likely to be different from that observed in plant material. Therefore, it must be evaluated at each locality.

The limit of the carbon-14 method based on standard counting techniques is about 35,000 years. Use of ultrasensitive mass spectrometers can more than double this limit. Dates are sometimes expressed in terms of years *before present* (B.P.), where the present is taken to be 1952. Historical dates are also referred to the Christian calendar by subtracting 1952 years from dates adjusted to that point of reference. Thus a date of 2000 years B.P. becomes 48 B.C. (before Christ), whereas a date of 1500 years B.P. is equivalent to 452 A.D. (*anno domini*).

Although the principles of the ^{14}C method of dating appear to be straightforward, many complications do, in fact, occur. These arise from three main sources:

1. Variations in the production rate of ^{14}C caused by fluctuations in the cosmic-ray flux and by changes in the magnetic field of the Earth, which interacts with the cosmic rays.
2. Variations of the ^{14}C content of plant materials and calcium carbonate caused by isotope fractionation and other local environmental effects.
3. Lowering of the ^{14}C content of the atmosphere by dilution with CO_2 derived by combustion of fossil fuel and increases caused by the explosion of nuclear devices and the operation of nuclear reactors.

The changes in the ^{14}C content of the atmosphere since the start of the industrial revolution around 1860 are so profound that the carbon-14 method of dating is no longer applicable to samples formed since that time.

b. The ^{10}Be and ^{26}Al Methods

Although the presence of ^{10}Be in rainwater has been known since 1955, its use in geological research was retarded by the technical difficulties in measuring its decay rate with radiation detec-

tors. The same is true for ^{26}Al , ^{36}Cl , and most of the other cosmogenic radionuclides listed in Table 16.2. All this changed dramatically in the late 1970s when ultrasensitive mass spectrometers were developed by means of which the abundances of these nuclides can now be measured accurately (Purser, et al., 1982).

Both ^{10}Be and ^{26}Al are produced in the atmosphere by highly energetic cosmic rays that fragment stable isotopes of O and N in *spallation reactions*. Both isotopes are rapidly removed from the atmosphere in the form of rain and snow. When they enter the oceans, both radionuclides are scavenged from the water by adsorption on inorganic and organic sediment particles that carry them down to the bottom of the oceans. After deposition, the concentrations (or disintegration rates) of ^{10}Be and ^{26}Al decrease with time as they decay with their characteristic half-lives (Table 16.2). The same is true when ^{10}Be and ^{26}Al are deposited on the continental ice sheets of Greenland and Antarctica. In addition, these and other radionuclides are produced in meteorites and in rocks on other planets exposed to cosmic rays. After meteorites fall to the Earth, the production of cosmogenic radionuclides virtually stops. This makes it possible to determine their *terrestrial residence ages* by measuring the remaining concentrations or activities of ^{10}Be and ^{26}Al .

The decay of ^{10}Be to stable ^{10}B by beta (β^-) decay obeys the law of radioactivity expressed by equation 16.9:

$$^{10}\text{Be} = ^{10}\text{Be}_0 e^{-\lambda t} \quad (16.51)$$

This equation can be solved for t based on measurements of ^{10}Be and $^{10}\text{Be}_0$. When samples are taken from a core of sediment or ice, t can be expressed in terms of depth (h) and the sedimentation rate (s) as in equation 16.42. When data points are plotted in coordinates of $\ln ^{10}\text{Be}$ and h , they define a straight line whose slope yields the sedimentation rate (s) and whose intercept is $\ln ^{10}\text{Be}_0$, as exemplified by equation 16.44.

In principle, the same suite of samples can be analyzed for both ^{10}Be and ^{26}Al , which act as independent chronometers. On the other hand, the two

radionuclides can be combined by taking the ratio of their concentrations or disintegration rates:

$$\frac{^{26}\text{Al}}{^{10}\text{Be}} = \left(\frac{^{26}\text{Al}}{^{10}\text{Be}} \right)_0 \frac{e^{-\lambda_{\text{Al}} t}}{e^{-\lambda_{\text{Be}} t}} \quad (16.52)$$

Equation 16.52 can be reduced to:

$$\frac{^{26}\text{Al}}{^{10}\text{Be}} = \left(\frac{^{26}\text{Al}}{^{10}\text{Be}} \right)_0 e^{-\lambda' t} \quad (16.53)$$

where

$$\begin{aligned} \lambda' &= \lambda_{\text{Al}} - \lambda_{\text{Be}} \\ &= (0.968 - 0.462) \times 10^{-6} \text{ a}^{-1} \\ &= 0.506 \times 10^{-6} \text{ a}^{-1} T_{1/2} = 1.37 \times 10^6 \text{ a} \end{aligned} \quad (16.54)$$

According to equation 16.53, the $^{26}\text{Al}/^{10}\text{Be}$ ratio decreases with depth in a sediment or ice core at a rate that depends on the rate of deposition. By substituting $t = h/s$ and by taking natural logarithms of equation 16.53 we obtain:

$$\ln \left(\frac{^{26}\text{Al}}{^{10}\text{Be}} \right) = \ln \left(\frac{^{26}\text{Al}}{^{10}\text{Be}} \right)_0 - \frac{\lambda' h}{s} \quad (16.55)$$

This is the equation of a straight line in coordinates of $\ln(^{26}\text{Al}/^{10}\text{Be})$ and h whose slope $m = -\lambda'/s$ and whose intercept is $\ln(^{26}\text{Al}/^{10}\text{Be})_0$. A limited number of data reviewed by Faure (1986) suggest that the initial $^{26}\text{Al}/^{10}\text{Be}$ ratio may be a constant equal to about 0.018 ± 0.010 .

c. Exposure Dating of Rock Surfaces

High-energy neutrons associated with cosmic rays produce radionuclides by nuclear spallation reactions involving atoms in rocks and soils exposed at the surface of the Earth (Lal, 1988). The products of these nuclear reactions include ^3He , ^{10}Be , ^{14}C , ^{21}Ne , ^{26}Al , and ^{36}Cl , which accumulate in the target rocks as a function of time. The decay constants and half-lives of the radionuclides are included in Table 16.2, except for ^3He and ^{21}Ne , which are stable.

Although nuclear spallation reactions occur in all rocks and minerals that are exposed to cosmic rays, quartz is an especially favorable target for the production of ^{10}Be and ^{26}Al because of its simple chemical composition (SiO_2), its resistance to mechanical and chemical weathering, and its

abundance in different kinds of rocks and because cosmogenic radionuclides deposited by meteoric precipitation can be eliminated by a chemical cleaning procedure (Lal and Arnold, 1985; Nishiizumi et al., 1986). In addition, in-situ-produced cosmogenic ^{36}Cl has been used successfully to determine exposure ages of lacustrine sediment and glacial deposits of Pleistocene age (Phillips et al., 1986; Phillips et al., 1990; Jannik et al., 1991).

The accumulation of a radionuclide formed by a nuclear reaction is governed by an equation derivable from the law of radioactivity (Faure, 1986, Section 4.4):

$$P = \frac{R}{\lambda}(1 - e^{-\lambda t}) \quad (16.56)$$

where P = number of product atoms per gram; R = production rate of the product nuclide in atoms per gram per year; λ = decay constant of the product in reciprocal years; and t = duration of the irradiation in years. The production rates of ^{10}Be and ^{26}Al in quartz at sea level and at geomagnetic latitude $>50^\circ$ have been determined by Nishiizumi et al. (1989) who reported rates of $6.0 \text{ atoms g}^{-1} \text{ y}^{-1}$ for ^{10}Be and $36.8 \text{ atoms g}^{-1} \text{ y}^{-1}$ for ^{26}Al . The production rates rise with increasing altitude above sea level because of decreased atmospheric shielding and also vary with geomagnetic latitude.

When equation 16.56 is used to determine the exposure age of quartz based on the measured concentrations of ^{10}Be or ^{26}Al , several assumptions must be made:

1. The quartz in the target rock did not contain residual cosmogenic radionuclides left over from a previous exposure.
2. The target rock was not shielded from cosmic rays by nearby topographic features.
3. The target rock was irradiated continuously since its initial exposure to cosmic rays.
4. The rate of erosion of the target rock was negligibly low.

These assumptions cannot be evaluated in most cases and therefore introduce uncertainty into the exposure dates calculated from equation 16.56.

The erosion rate of the target rock can be taken into consideration provided it remained constant during the exposure. The relevant equation was derived by Lal (1991):

$$P = \frac{R}{\lambda + \rho\varepsilon/\Lambda}(1 - e^{-(\lambda + \rho\varepsilon/\Lambda)t}) \quad (16.57)$$

where ε = erosion rate of the target in centimeters per year; ρ = density of the target mineral in grams per cubic centimeters; Λ = absorption mean free path of neutrons in the target rock. Nishiizumi et al. (1991) used a value of 165 g/cm^2 for this parameter. Note that the decay constant of the product nuclide in equation 16.57 has been increased by the term $\rho\varepsilon/\Lambda$ and that the apparent half-life has therefore been shortened. In other words, the effect of erosion is to make the product nuclide appear to decay more rapidly than it actually does. Therefore, a longer exposure time is required to achieve the observed concentration of the cosmogenic radionuclide, meaning that the exposure age calculated from equation 16.57 increases with increasing erosion rate.

Equation 16.57 can be used to calculate a minimum exposure date and a maximum erosion rate based on the measured concentrations of ^{10}Be or ^{26}Al in a sample of quartz that was exposed to cosmic rays. If the erosion rate is assumed to be negligibly small, equation 16.57 approaches the form of equation 16.56 and yields a minimum exposure age. However, at higher erosion rates the corresponding exposure ages calculated from equation 16.57 rapidly approach infinity as the erosion rate reaches a maximum. Therefore, the cosmic-ray-exposure date calculated from equation 16.56 is an *underestimate* of the true exposure age.

The discrepancy between the calculated minimum date and the true exposure age rises with increasing values of the erosion rate of the target and with the duration of the exposure. In addition, the sensitivity of this method of dating decreases as the concentration of the product nuclide approaches equilibrium when the rate of production of the radionuclide is equal to its rate of decay. This condition is achieved when the irradiation time is equal to about four half-lives of the

product nuclide. Since ^{10}Be has a half-life of 1.5×10^6 years, equilibrium is reached in about 6 million years, assuming that the erosion rate of the target is negligible. This means that a calculated exposure date of that magnitude may actually be older by amounts of time that cannot be specified. However, exposure ages of less than about one million years, calculated from in-situ-produced ^{10}Be in quartz, are reliable, provided that the erosion rates are low (i.e. less than about 10^{-5} cm/year) and that the other assumptions listed above are satisfied. Therefore, the most desirable material for exposure dating are quartz crystals, vein quartz, chert, quartz concretions or agate geodes, quartzite, or well-cemented quartz arenites.

Age determinations based on in-situ-produced ^{10}Be in quartz have been reported by Nishiizumi (1991a,b), Brook et al. (1993), Gosse et al. (1995).

16.5 Summary

The discovery of the radioactivity of the salts of U in 1896 has had a profound effect on geology because radioactivity liberates heat and because it provides clocks that tell geologic time. These geochronometers are based on the slow rate of decay of long-lived radioactive isotopes of K, Rb,

Sm, La, Lu, Re, U, and Th. The interpretation of the measurements is based on the law of radioactivity, which yields information not only about time but also about the past histories of the daughter elements within the Earth. Therefore, isotopic methods of dating based on long-lived radioactive isotopes provide information not only about the history of the Earth but also about the origin of igneous rocks and the chemical differentiation of the crust and mantle.

The short-lived radioactive daughters of ^{238}U and ^{235}U enable us to determine sedimentation rates and ages of minerals that formed within 1 million years of the present. The study of geologic processes of the recent past is especially important because it contributes to our understanding of the terrestrial environment, including variations in climate.

Cosmogenic radionuclides are becoming increasingly important for dating events on the surface of the Earth within the past million years. The nuclides include not only the well-known ^{14}C but also ^{10}Be , ^{26}Al , and ^{36}Cl produced by exposure of rocks and minerals to cosmic rays. The increasing emphasis on dating events that occurred in the recent past reflects the growing conviction that we must understand the recent history of the Earth in order to predict its future.

Problems*

*Problems 2–10 are based on data from Baadsgaard et al. (1976), and Problems 11 and 12 are from Baadsgaard (1973) concerning the geologic history of the Amitsoq gneiss of Greenland, which is one of the oldest rocks known on Earth.

1. Calculate the atomic $^{87}\text{Rb}/^{86}\text{Sr}$ ratio of a rock sample containing $\text{Rb} = 125 \mu\text{g/g}$ and $\text{Sr} = 275 \mu\text{g/g}$. Atomic weights: $\text{Rb} = 85.46$, $\text{Sr} = 87.62$; isotopic abundances: $^{87}\text{Rb} = 27.83\%$, $^{86}\text{Sr} = 9.90\%$.

(Answer: 1.31)

2. Calculate a K–Ar date for a sample of biotite from the Amitsoq gneiss containing $\text{K}_2\text{O} = 9.00\%$, $^{40}\text{Ar} = 2.424 \mu\text{g/g}$. Decay constants: $\lambda_e = 0.581 \times 10^{-10} \text{ a}^{-1}$, $\lambda = 5.543 \times 10^{-10} \text{ a}^{-1}$; isotopic abundance of $^{40}\text{K} = 0.01167\%$; atomic weights: $\text{K} = 39.098$, $\text{O} =$

15.999 , $^{40}\text{Ar} = 39.96$.

(Answer: $2308 \times 10^6 \text{ a}$)

3. Calculate a Rb–Sr date for the same sample of biotite from the Amitsoq gneiss given that $^{87}\text{Rb}/^{86}\text{Sr} = 107.1$, $^{87}\text{Sr}/^{86}\text{Sr} = 3.093$. Decay constant: $\lambda = 1.42 \times 10^{-11} \text{ a}^{-1}$; $(^{87}\text{Sr}/^{86}\text{Sr})_0 = 0.7030$.

(Answer: $1554 \times 10^6 \text{ a}$)

4. Calculate the slope and intercept of a Rb–Sr isochron by the method of least squares using data for micas from the Amitsoq gneiss.

Samples	$^{87}\text{Rb}/^{86}\text{Sr}$	$^{87}\text{Sr}/^{86}\text{Sr}$
1a	763	17.337
1b	41.52	1.739
2	107.12	3.093
3	166.7	4.543
4	138.7	3.858
5	330.7	8.057
6	82.7	2.741

Use the slope to calculate a date for these samples ($\lambda = 1.42 \times 10^{-11} \text{ a}^{-1}$).

(Answer: intercept = 0.8806, slope = 0.021 59, $t = 1504 \times 10^6 \text{ a}$)

5. Calculate the Rb–Sr isochron age and initial $^{87}\text{Sr}/^{86}\text{Sr}$ ratio of a suite of whole-rock samples from the Amitsoq gneiss near Godthaab, Greenland, based on the following data ($\lambda = 1.42 \times 10^{-11} \text{ a}^{-1}$).

Sample	$^{87}\text{Rb}/^{86}\text{Sr}$	$^{87}\text{Sr}/^{86}\text{Sr}$
1	2.098	0.8245
2	0.198	0.7096
3	1.173	0.7668
4	2.033	0.8191
5	1.364	0.7791
6	0.319	0.7163

[Answer: $(^{87}\text{Sr}/^{86}\text{Sr})_0 = 0.6971$, slope = 0.0060 22, $t = 4119 \times 10^6 \text{ a}$. For a discussion of these results consult the paper by Baadsgaard et al. (1976)]

6. Calculate a Rb–Sr date from the following analyses of a whole rock and its constituent minerals ($\lambda = 1.42 \times 10^{-11} \text{ a}^{-1}$).

Sample	$^{87}\text{Rb}/^{86}\text{Sr}$	$^{87}\text{Sr}/^{86}\text{Sr}$
Whole rock	1.173	0.7668
Biotite	116.7	4.543
K-feldspar	0.6437	0.7822
Plagioclase	0.0633	0.7344

Note that the whole rock is sample 3 in Problem 5 and the biotite is sample 3 in Problem 4.

[Answer: $(^{87}\text{Sr}/^{86}\text{Sr})_0 = 0.7468$, slope = 0.022 77, $t = 1586 \times 10^6 \text{ a}$]

7. Plot the whole-rock isochron from Problem 5 and the mineral isochron from Problem 6 and use the diagram to interpret the dates obtained in Problem 5 and 6.

8. Calculate U–Pb and Th–Pb dates for a sample of sphene from the Amitsoq gneiss in Greenland. $^{206}\text{Pb}/^{204}\text{Pb} = 53.90$, $^{207}\text{Pb}/^{204}\text{Pb} = 20.76$, $^{208}\text{Pb}/^{204}\text{Pb} = 36.94$, $^{238}\text{U}/^{204}\text{Pb} = 88.746$, $^{238}\text{U}/^{235}\text{U} = 137.88$, $^{232}\text{Th}/^{204}\text{Pb} = 55.38$, $(^{206}\text{Pb}/^{204}\text{Pb})_0 = 12.97$, $(^{207}\text{Pb}/^{204}\text{Pb})_0 = 14.17$, $(^{208}\text{Pb}/^{204}\text{Pb})_0 = 33.90$, $\lambda_1 = 1.55125 \times 10^{-10} \text{ a}^{-1}$, $\lambda_2 = 9.8485 \times 10^{-10} \text{ a}^{-1}$, $\lambda_3 = 4.9475 \times 10^{-11} \text{ a}^{-1}$.

(Answer: $t_{206} = 2445 \times 10^6 \text{ a}$, $t_{207} = 2456 \times 10^6 \text{ a}$, $t_{208} = 1080 \times 10^6 \text{ a}$)

9. Calculate the U/Th concentration ratio of the sphene in Problem 8 given that $^{238}\text{U}/^{204}\text{Pb} = 88.746$ and $^{232}\text{Th}/^{204}\text{Pb} = 55.38$. Atomic weights: U = 238.029, Th = 232.038. Abundance of $^{238}\text{U} = 99.28\%$.

(Answer: U/Th = 1.655)

10. Calculate the 207/206 date of sphene from the Amitsoq gneiss given that $^{206}\text{Pb}/^{204}\text{Pb} = 140.8$, $^{207}\text{Pb}/^{204}\text{Pb} = 35.36$, $(^{206}\text{Pb}/^{204}\text{Pb})_0 = 12.97$, $(^{207}\text{Pb}/^{204}\text{Pb})_0 = 14.17$.

(Answer: $t_{207/206} = 2512 \times 10^6 \text{ a}$)

11. Plot a concordia curve for values of t from 3.0×10^9 to $3.7 \times 10^9 \text{ a}$.

12. Determine the age of zircons from the Amitsoq gneiss by plotting the data given below on the diagram constructed above (Baadsgaard, 1973).

Sample	$^{206}\text{Pb}^*/^{238}\text{U}$	$^{207}\text{Pb}^*/^{235}\text{U}$
1	0.732	32.25
2	0.715	32.25
3	0.714	31.75
4	0.673	29.75
5	0.643	27.50
6	0.625	27.00
7	0.600	25.50
8	0.595	25.60
9	0.500	20.00

$\lambda_1 = 1.55125 \times 10^{-10} \text{ a}^{-1}$, $\lambda_2 = 9.8485 \times 10^{-10} \text{ a}^{-1}$.

(Answer: discordia intercept = 0.1324, slope = 0.018 30, intercept with concordia: $^{207}\text{Pb}^*/^{235}\text{U} = 33.74$, $^{206}\text{Pb}^*/^{238}\text{U} = 0.754$, $t = 3622 \times 10^6 \text{ a}$)

13. Review the answers to Problems 2–6, 8, 10, and 12, all of which pertain to the Amitsoq gneisses.

(a) Why is the K–Ar date of biotite (Problem 2) older than the Rb–Sr date of the same biotite (Problem 3)?

(b) What is the significance of the fact that micas from different whole-rock samples of the Amitsoq gneisses all fit the same isochron (Problem 4)?

(c) What is the meaning of the whole-rock isochron date of the Amitsoq gneisses (Problem 5)?

(d) Why does the whole-rock sample (3) in Problem 5 also fit the isochron in Problem 6?

(e) What is the significance of the U–Pb dates of sphene from the Amitsoq gneisses in Problems 8 and 10?

(f) Suggest a reason for the low Th–Pb date (Problem 8) and the low U/Th ratio (Problem 9) of the sphenes from the Amitsoq gneisses.

(g) What is the significance of the concordia date of the zircons from the Amitsoq gneisses?

(h) What is the significance of the date of the lower point of intersection of discordia with concordia at $\sim 950 \times 10^6$ a? (Baadsgaard, 1973.)

(i) Why is the whole-rock Rb–Sr isochron date of the Amitsoq gneisses older than the concordia U–Pb date of the zircons?

14. Calculate the average growth rate of a manganese nodule from the following information.

Depth, mm	$(^{230}\text{Th}/^{232}\text{Th})_A$, dpm/g
0–0.3	43.7
0.3–0.5	26.2
0.5–0.85	10.4

$$\lambda = 9.217 \times 10^{-6} \text{ a}^{-1}.$$

(Answer: 3.2 mm/ 10^6 yr)

15. A sample of wood from an archeological site was found to be 6367 years old when it was analyzed in 1986. What is its age (a) in years B.P.? (b) in reference to the Christian calendar?

[Answer: (a) 6333 years B.P.; (b) 4381 B.C.]

References

General

- BAADSGAARD, H., 1973. U–Th–Pb dates on zircons from the early Precambrian Amitsoq gneisses, Godthaab District, West Greenland. *Earth Planet. Sci. Lett.*, **19**:22–28.
- BAADSGAARD, H., R. ST. J. LAMBERT, and J. KRUPICKA, 1976. Mineral isotopic age relationships in the polymetamorphic Amitsoq gneisses, Godthaab District, West Greenland. *Geochim. Cosmochim. Acta*, **40**:513–527.
- BATEMAN, H., 1910. Solution of a system of differential equations occurring in the theory of radioactive transformations. *Proc. Cambridge Phil. Soc.*, **15**:423.
- BECQUEREL, H., 1896. Sur les radiations invisibles émises par les sels d'uranium. *Compt. Rend. Acad. Sci. Paris*, **122**:689.
- BROECKER, W. S., D. L. THURBER, J. GODDARD, T.-L. KU, R. K. MATHEWS, and K. J. MESOLELLA, 1968. Milankovitch hypothesis supported by precise dating of coral reefs and deep-sea sediment. *Science*, **159**:287–300.
- BURCHFIELD, J. D., 1975. *Lord Kelvin and the Age of the Earth*. Science History Publications, New York, 260 pp.
- CREASER, R. A., D. A. PAPANASTASSIOU, and G. J. WASSERBURG, 1991. Negative thermal ion mass spectrometry of osmium, rhenium, and iridium. *Geochim. Cosmochim. Acta*, **55**:397–401.
- CURIE, P., and A. LABORDE, 1903. Sur la chaleur dégagée spontanément par les sels de radium. *Compt. Rend. Acad. Sci. Paris*, **136**:673–675.
- DALRYMPLE, G. B., and M. A. LANPHERE, 1969. *Potassium–Argon Dating*. W. H. Freeman, New York, 258 pp.
- DICKIN, A. P., 1995. *Radiogenic Isotope Geology*. Cambridge University Press, New York, 425 pp.
- FAURE, G., and J. L. POWELL, 1972. *Strontium Isotope Geology*. Springer-Verlag, New York, 188 pp.
- FAURE, G., 1986. *Principles of Isotope Geology*, 2nd ed. Wiley, New York, 589 pp.
- HOLMES, A., 1911. The association of lead with uranium in rock-minerals, its application to the measurement of geologic time. *Proc. Roy. Soc. London, Ser. A*, **85**:248–256.
- HOLMES, A., 1913. *The Age of the Earth*. Harper and Brothers, London, 194 pp.
- JOLY, J. J., 1909. *Radioactivity and Geology*. Archibald Constable, London, 287 pp.

- KROGH, T. E., 1982a. Improved accuracy of U-Pb zircon dating by selection of more concordant fractions using a high gradient magnetic separation technique. *Geochim. Cosmochim. Acta*, **46**:631-635.
- KROGH, T. E., 1982b. Improved accuracy of U-Pb zircon ages by the creation of more concordant systems using an air abrasion technique. *Geochim. Cosmochim. Acta*, **46**:637-649.
- ROMER, A., 1971. *Radiachemistry and the Discovery of Isotopes*. Dover, New York, 261 pp.
- RUTHERFORD, E., and F. SODDY, 1902. The cause and nature of radioactivity. Part I. *Phil. Mag., Ser. 6*, **4**:370-396.
- RUTHERFORD, E., 1906. *Radioactive Transformations*. Charles Scribner's Sons, New York, 287 pp.
- STACEY, J. S., and J. D. KRAMERS, 1975. Approximation of terrestrial lead isotope evolution by a two-stage model. *Earth Planet. Sci. Lett.*, **26**:207-221.
- STEIGER, R. H., and E. JAGER, 1977. Subcommission on geochronology: Convention on the use of decay constants in geochronology and cosmochronology. *Earth Planet. Sci. Lett.*, **36**:359-362.
- THOMSON, W. (Lord Kelvin), 1899. The age of the Earth as an abode fitted for life. *Phil. Mag.*, (5), **47**:66-90.
- WETHERILL, G. W., 1956. Discordant uranium-lead ages. *Trans. Amer. Geophys. Union*, **37**:320-326.
- ### Cosmogenic
- ANDERSON, E. C., and W. F. LIBBY, 1951. Worldwide distribution of natural radiocarbon. *Phys. Rev.*, **81**:64-69.
- ARNOLD, J. R., and W. F. LIBBY, 1949. Age determinations by radiocarbon content: Checks with samples of known age. *Science*, **110**:678-680.
- BROOK, E. J., M. D. KURZ, R. P. ACKERT, JR., G. H. DENTON, E. T. BROWN, G. M. RAISBECK, and F. YIOU, 1993. Chronology of Taylor Glacier advances in Arena Valley, Antarctica, using in situ cosmogenic ^3He and ^{10}Be . *Quaternary Research*, **39**:11-23.
- BROOK, E. J., E. T. BROWN, M. D. KURZ, R. P. ACKERT JR., G. M. RAISBECK, and F. YIOU, 1995. Constraints on age, erosion, and uplift of Neogene glacial deposits in the Transantarctic Mountains determined from in situ cosmogenic ^{10}Be and ^{26}Al . *Geology*, **23**(12):1063-1066.
- GOSSE, J. C., E. B. EVENSON, J. KLEIN, B. LAWN, R. MIDDLETON, 1995. Precise cosmogenic ^{10}Be measurements in western North America: Support for a global Younger Dryas cooling event. *Geology*, **23**(10):877-880.
- IVY-OCHS, S., C. SCHLÜCHTER, P. W. KUBIK, B. DITTRICH-HANNEN, and J. BEER, 1995. Minimum ^{10}Be exposure ages of early Pliocene for the Table Mountain plateau and the Sirius Group at Mount Fleming, Dry Valleys, Antarctica. *Geology*, **23**(11):1007-1010.
- JANNIK, N. O., F. M. PHILLIPS, G. I. SMITH, D. ELMORE, 1991. A ^{36}Cl chronology of lacustrine sedimentation in the Pleistocene Owens River system. *Geol. Soc. Amer. Bull.*, **103**(9):1146-1159.
- LAL, D., 1988. In situ-produced cosmogenic isotopes in terrestrial rocks. *Ann. Rev. Earth Planet. Sci.*, **16**:355-388.
- LAL, D., 1991. Cosmic ray labeling of erosion surfaces: in situ nuclide production rates and erosion models. *Earth Planet. Sci. Letters*, **104**:424-439.
- LAL, D., and J. R. ARNOLD, 1985. Tracing quartz through the environment. *Proc. Indian Acad. Sci. (Earth Planet. Sci.)*, **94**(1):1-5.
- LIBBY, W. F., 1952. *Radiocarbon Dating*. University of Chicago Press, 175 pp.
- MOOK, W. G. (Ed.), 1980. *The Principles and Applications of Radiocarbon Dating*. Elsevier, Amsterdam.
- NISHIZUMI, K., D. LAL, J. KLEIN, R. MIDDLETON, and J. R. ARNOLD, 1986. Production of ^{10}Be and ^{26}Al by cosmic rays in terrestrial quartz in situ and implications for erosion rates. *Nature*, **319**:134-136.
- NISHIZUMI, K., E. L. WINTERER, C. P. KOHL, and J. KLEIN, 1989. Cosmic ray production rates of ^{10}Be and ^{26}Al in quartz from glacially polished rocks. *J. Geophys. Res.*, **94**(B12):17,907-17,915.
- NISHIZUMI, K., C. P. KOHL, J. R. ARNOLDS, J. KLEIN, D. FINK and R. MIDDLETON, 1991a. Cosmic ray produced ^{10}Be and ^{26}Al in Antarctic rocks: exposure and erosion history. *Earth Planet. Sci. Letters*, **104**:440-454.
- NISHIZUMI, K., C. P. KOHL, E. M. SHOEMAKER, J. R. ARNOLD, J. KLEIN, D. FINK, and R. MIDDLETON, 1991b. In situ ^{10}Be - ^{26}Al exposure ages at Meteor Crater, Arizona. *Geochim. Cosmochim. Acta*, **55**:2699-2703.
- PHILLIPS, F. M., B. D. LEAVY, N. O. JANNIK, D. ELMORE, and P. W. KUBIK, 1986. The accumulation of cosmogenic chlorine-36 in rocks: A method for surface exposure dating. *Science*, **231**:41-42.
- PHILLIPS, F. M., M. G. ZREDA, S. S. SMITH, D. ELMORE, P. W. KUBIK, and P. SHARMA, 1990. A cosmogenic chlorine-36 chronology of glacial deposits at Bloody Canyon, eastern Sierra Nevada, California. *Science*, **248**:1529-1532.
- PURSER, K. H., C. J. RUSSO, R. B. LIEBERT, H. E. GOVE, D. ELMORE, R. FERRARO, A. E. LITHERLAND, R. P. BEUKENS, K. H. CHANG, L. R. KILIUS, and H. W. LEE, 1982. The application of electrostatic tandems to ultrasensitive mass spectrometry and nuclear dating. In L. A. Currie (Ed.), *Nuclear and Chemical Dating Techniques*, 45-74. *Amer. Chem. Soc., Symp. Ser.*, vol. 176, 516 pp.

Isotope Fractionation

Several elements of low atomic number are composed of two or more stable isotopes. Therefore, when these elements react to form compounds, the molecules that are produced differ from each other by the particular isotopes they contain. For example, there are nine different isotopic varieties of H_2O and ten isotopic varieties of CO_2 whose molecular weights differ significantly from each other. The differences in the molecular weights affect the way these molecules respond to certain kinds of physical processes that are mass dependent. Moreover, the strengths of covalent bonds in molecules or ions depend on the masses of the atoms. Therefore, certain isotopic varieties of a particular molecule or complex ion are more stable than others because the bonds they contain are stronger than those of other isotopic varieties of the same molecule. The study of these subtle effects has become very important in the Earth Sciences because they enhance our understanding of geochemical processes occurring on the surface of the Earth.

17.1 Principles of Isotope Fractionation

The isotope compositions of certain elements of low atomic number are variable and depend on the compound in which the element resides as well as on the origin of that compound. For example, the isotopic composition of oxygen in rainwater varies widely depending on the latitude and other geographic and climatic factors of the collecting site. Similarly, seawater, marine mollusk shells, and quartz grains on a beach all contain oxygen of different isotope composi-

tions. The observed differences in the isotope compositions of oxygen and other elements of low atomic number are caused by processes referred to as *isotope fractionation*.

The magnitude of isotope fractionation depends on differences in the masses of the isotopes, on the temperature of formation of the compounds in which they occur, on the character of the chemical bonds they form, and on other factors. The elements whose isotopes are commonly fractionated include H, C, N, O, and S (see Table 17.1). In addition, isotope fractionation has been observed in Li, B, and Si. However, isotope fractionation is not possible for Be, F, Na, Al, and P because these elements have only one stable isotope.

Variations in the isotopic compositions of H, C, N, O, and S are important in the study of natural processes because they are among the most abundant elements in the crust of the Earth, because they form a wide variety of solid, liquid, and gaseous compounds, and because they occur in the crust and mantle of the Earth, as well as in its biosphere, hydrosphere, and atmosphere. Therefore, these elements participate in virtually all geochemical processes that take place on the surface of the Earth. As a result of the fractionation of their isotopes by these processes, the elements preserve a record of the environmental conditions under which their compounds formed. The growing importance of isotope fractionation in the study of sedimentary rocks and natural environments is reflected in books by Faure (1986), Hoefs (1996), Arthur et al. (1983), Fritz and Fontes (1980a, 1986), Rundel et al. (1988), and Ding et al. (1996).

Isotope fractionation is caused by isotope exchange reactions and by mass-dependent

Table 17.1 Abundances and Masses of Stable Isotopes of Elements Subject to Isotope Fractionation

Element	Stable isotopes	Mass, amu	Average abundance, %	Mass ^a difference, %
Hydrogen	¹ H	1.007825	99.985	99.8
	² H (D) ^b	2.0140	0.015	
Carbon	¹² C	12.000000	98.90	8.36
	¹³ C	13.003355	1.10	
Nitrogen	¹⁴ N	14.003074	99.63	7.12
	¹⁵ N	15.000108	0.37	
Oxygen	¹⁶ O	15.994915	99.762	12.5
	¹⁷ O	16.999131	0.038	
	¹⁸ O	17.999160	0.200	
Sulfur	³² S	31.972070	95.02	6.24 (³⁴ S - ³² S)
	³³ S	32.971456	0.75	
	³⁴ S	33.967866	4.21	
	³⁶ S	35.967080	0.02	

^aThe mass difference is expressed as $(m_H - m_L)/m_L \times 10^2$, where m_H and m_L are the masses of molecules containing the heavy and light isotopes, respectively.

^bDeuterium symbolized by D.

SOURCE: Weast et al. (1986).

differences in the rates of certain chemical reactions and physical processes. For example, an important exchange reaction involving two isotopic molecules of CO₂ and H₂O is:



In this reaction ¹⁶O and ¹⁸O are exchanged between the two molecules and the factor of $\frac{1}{2}$ is needed to reduce the number of O atoms in CO₂ to one. Isotope exchange reactions progress from some initial state toward equilibrium much like ordinary chemical reactions. At equilibrium the Law of Mass Action applies:

$$\frac{(\text{C}^{18}\text{O}_2)^{1/2}(\text{H}_2^{16}\text{O})}{(\text{C}^{16}\text{O}_2)^{1/2}(\text{H}_2^{18}\text{O})} = K \quad (17.2)$$

where K is the equilibrium constant. The amounts of reactants and products are expressed in terms of molar concentrations rather than activities because the activity coefficients of isotopic varieties of a molecule cancel in equations like 17.2. If

CO₂ and H₂O molecules accept ¹⁶O and ¹⁸O without discrimination, then $K = 1.00$ and no isotope fractionation takes place. In fact, the equilibrium constant of reaction 17.1 at 25°C is $K = 1.0412$, which implies that the molecules do favor one of the O isotopes over the other.

The standard free energy change of isotope exchange reactions is of the order of -0.020 kcal, which is quite low compared to free energy changes of chemical reactions (Valley et al., 1986). Nevertheless, isotope exchange reactions can achieve equilibrium in spite of the low driving force, and the existence of isotopic equilibrium generally implies that the chemical reactions are also in equilibrium.

The equilibrium constant of the isotope exchange reaction represented by equation 17.1 differs from unity at 25°C because ¹⁸O forms a *stronger* covalent bond with C than ¹⁶O does. The explanation for this fact arises from the quantum mechanical treatment of the energy of a diatomic

molecule. The strength of the covalent bond in such a molecule arises from the *reduction* of the energy of the molecule as the two atoms approach each other. At room temperature, most molecules of a gas are in the ground vibrational state (quantum number $n = 1$) and their vibrational energy (E) is:

$$E = \frac{1}{2}h\nu \quad (17.3)$$

where h is Planck's constant and ν is the frequency of vibration. When the heavy isotope of an element replaces the light isotope, the energy of the molecule *decreases* because the increase in the mass of the molecule decreases its vibrational frequency. Consequently, the heavy isotope forms a stronger bond than the light isotope and molecules containing the heavy isotope of an element are more stable than those containing the light isotope. The *general rule* regarding isotope exchange reactions stated by Bigeleisen (1965) is "The heavy isotope goes preferentially to the chemical compound in which the element is bound most strongly."

The second cause for isotope fractionation is related to the differences in the masses of isotopes and of the molecules in which they occur. For example, the data in Table 17.1 indicate that deuterium (${}^2_1\text{H}$) is 99.8% heavier than protium (${}^1_1\text{H}$). The mass differences for the isotopes of the other elements are less than that, but still average $\sim 8.6\%$ and decrease with increasing atomic number of the elements. The molecular weights of *compounds* also vary depending on the isotopes they contain. For example, the molecular weight of ${}^2\text{H}_2{}^{16}\text{O}$ (or $\text{D}_2{}^{16}\text{O}$) is 11.1% greater than that of ${}^1\text{H}_2{}^{16}\text{O}$ (or $\text{H}_2{}^{16}\text{O}$). Similarly, $\text{H}_2{}^{18}\text{O}$ is also $\sim 11.1\%$ heavier than $\text{H}_2{}^{16}\text{O}$. Such differences in the molecular weights of compounds affect the rates of certain physical and chemical processes like diffusion, evaporation, and dissociation.

The fractionation of isotopes during diffusion and evaporation is caused by differences in the velocities of isotopic molecules of the same compound. According to statistical mechanics, all molecules of an ideal gas have the same kinetic energy depending on the temperature. Since the

kinetic energy is equal to $\frac{1}{2}mv^2$, the ratio of the velocities (v) of two isotopic molecules of mass (m) of the same compound is:

$$\frac{v_L}{v_H} = \left(\frac{m_H}{m_L}\right)^{1/2} \quad (17.4)$$

where the subscripts L and H identify the molecule containing the *light* and *heavy* isotopes, respectively. In the case of the carbon monoxide molecules ${}^{12}\text{C}{}^{16}\text{O}$ and ${}^{13}\text{C}{}^{16}\text{O}$, the ratio of the velocities is:

$$\frac{v_L}{v_H} = \left(\frac{28.998\,27}{27.994\,915}\right)^{1/2} = 1.0177 \quad (17.5)$$

This means that the velocity of ${}^{12}\text{C}{}^{16}\text{O}$ is 1.0177 times greater than that of ${}^{13}\text{C}{}^{16}\text{O}$ regardless of the temperature.

Velocity differences among isotopic molecules of the same compound contribute to isotope fractionation in a number of ways. For example, molecules containing the light isotope of an element may diffuse out of a system more rapidly than the molecules containing the heavy isotope, thereby enriching the system in the heavy isotope. Therefore, during evaporation of water, $\text{H}_2{}^{16}\text{O}$ molecules escape into the vapor phase more rapidly than $\text{H}_2{}^{18}\text{O}$ and thus enrich the vapor in ${}^{16}\text{O}$ leaving the remaining water enriched in ${}^{18}\text{O}$.

Differences in bond strengths among isotopic molecules of the same compound affect the kinetics of reactions in which they participate. Since the lighter of two isotopes forms *weaker* bonds than the heavier isotope, molecules containing the lighter isotope react more rapidly and are concentrated preferentially into the products of the reaction. Therefore, unidirectional chemical reactions can be very effective in fractionating isotopes. Such reactions occur in open systems where the products can escape and equilibrium is unattainable. For example, certain bacteria (*Desulfovibrio desulfuricans*) reduce sulfate ions to H_2S^0 or HS^- , which may be removed from the reaction by precipitation of metal sulfides. The sulfides produced by this process are enriched in ${}^{32}\text{S}$ relative to ${}^{34}\text{S}$.

because the ^{32}S —O bonds of the sulfate ions are more easily broken by the bacteria than the ^{34}S —O bonds.

The equilibrium constants of isotope exchange reactions in the gas phase can be calculated from so-called partition functions, which are defined in statistical mechanics. However, the fractionation of isotopes by reactions involving solids or liquids must be studied under controlled conditions in the laboratory or by analysis of natural systems believed to have achieved isotopic equilibrium. The magnitude of equilibrium constants of isotope exchange reactions depends primarily on the temperature and approaches 1.00 as the temperature rises to high values because at high temperature the small differences in the strengths of bonds formed by different isotopes of an element in a molecule are not as effective in causing isotope fractionation as they are at low temperature.

17.2 Mathematical Relations

The equilibrium constant of an isotope exchange reaction involving the replacement of only *one* atom in a molecule by two different isotopes of an element is identical to the *isotope fractionation factor*, which is defined by:

$$\alpha_b^a = \frac{R_a}{R_b} \quad (17.6)$$

where R_a is the ratio of the heavy to the light isotope of an element in a molecule or phase *a*, R_b is the same in molecule or phase *b*, and the reaction is in equilibrium at a specific temperature. For example, when water is in equilibrium with its vapor at 25°C, the fractionation of O isotopes in water molecules is expressed by:

$$\alpha_v^l = \frac{R_l}{R_v} = 1.0092 \quad (17.7)$$

where $R_l = ^{18}\text{O}/^{16}\text{O}$ in the liquid phase and $R_v = ^{18}\text{O}/^{16}\text{O}$ in the vapor phase. The corresponding fractionation factor for H between water and

its vapor is defined in terms of the D/H ratios and has a value of 1.074 at 25°C.

The data in Table 17.1 indicate that the $^{18}\text{O}/^{16}\text{O}$ ratio of *average* terrestrial O is $0.200/99.762 = 0.00200$ and that the D/H ratio is 0.00015. Because it is difficult to measure small differences in such small ratios, the isotope compositions of H, C, N, and S are expressed as the per mille difference of the isotope ratios of a sample (spl) and a standard (std). For O we have:

$$\delta^{18}\text{O} = \left(\frac{R_{\text{spl}} - R_{\text{std}}}{R_{\text{std}}} \right) \times 10^3 \text{‰} \quad (17.8)$$

where R is the ratio of the heavy to the light isotope, that is, $R = ^{18}\text{O}/^{16}\text{O}$. Similarly for C:

$$\delta^{13}\text{C} = \left(\frac{R_{\text{spl}} - R_{\text{std}}}{R_{\text{std}}} \right) \times 10^3 \text{‰} \quad (17.9)$$

where $R = ^{13}\text{C}/^{12}\text{C}$. The standard for O is *standard mean ocean water* (SMOW), whereas that for C is the calcite of a belemnite from the Peedee Formation in South Carolina known as PDB. SMOW also serves as the standard for H, whereas the isotope composition of O in carbonate rocks and minerals is commonly expressed relative to PDB. The relationship between $\delta^{18}\text{O}$ (SMOW) and $\delta^{18}\text{O}$ (PDB) is:

$$\delta^{18}\text{O} (\text{SMOW}) = 1.03086 \delta^{18}\text{O} (\text{PDB}) + 30.86 \quad (17.10)$$

The standard for N is N_2 of the atmosphere, whereas the isotopic composition of S is expressed relative to the S in troilite (FeS) of the Canyon Diablo meteorite from Meteor Crater, Arizona. In all cases, a positive δ -value indicates that the sample is enriched in the heavy isotope relative to the standard, whereas a negative δ -value indicates depletion in the heavy isotope. The δ -values are determined on specially designed mass spectrometers using gases prepared from the samples. The practical problems associated with the determination of δ -values are discussed by Hoefs (1996).

The relationship between the fractionation factor α_b^a for an element whose isotopes are dis-

tributed between molecules or phases a and b and the corresponding δ -values can be derived from their definitions. Since $\alpha_b^a = R_a/R_b$, we use equation 17.8 to express R_a and R_b in terms of δ -values:

$$R_a = \frac{\delta_a R_{\text{std}}}{10^3} + R_{\text{std}} = \frac{R_{\text{std}}(\delta_a + 10^3)}{10^3} \quad (17.11)$$

Therefore,

$$\alpha_b^a = \frac{R_{\text{std}}(\delta_a + 10^3)10^3}{10^3 R_{\text{std}}(\delta_b + 10^3)} = \frac{\delta_a + 10^3}{\delta_b + 10^3} \quad (17.12)$$

Equation 17.12 enables us to calculate either α_b^a , δ_a , or δ_b when any two of the three variables are known. For example, we can use equation 17.12 to calculate the $\delta^{18}\text{O}$ value of water vapor in the air over the oceans at 25°C, assuming that isotope equilibrium is maintained. Since $\alpha_v^l(\text{O}) = 1.0092$ and $\delta^{18}\text{O}_l = 0.0\text{‰}$ relative to SMOW:

$$\delta^{18}\text{O}_v = \frac{\delta^{18}\text{O}_l + 10^3}{\alpha_v^l} - 10^3 \quad (17.13)$$

$$\delta^{18}\text{O}_v = \frac{0 + 10^3}{1.0092} - 10^3 = -9.1\text{‰} \quad (17.14)$$

The minus sign indicates that water vapor is depleted in ^{18}O relative to SMOW. We can also calculate δD of water vapor in equilibrium with seawater at 25°C. Since $\alpha_v^l(\text{D}) = 1.074$ and $\delta\text{D}_l = 0.0\text{‰}$, $\delta\text{D}_v = -68.9\text{‰}$, which means that the vapor is depleted in deuterium by 68.9‰ or 6.89%. In reality, water vapor over the oceans is even *more* depleted in ^{18}O and D than predicted by these calculations because the air over the oceans is undersaturated with respect to water vapor, which means that seawater is not in equilibrium with its vapor. As a result, the rate of evaporation exceeds the rate of condensation and kinetic isotope fractionation takes place.

The temperature variation of isotope fractionation factors for specific exchange reactions is usually determined experimentally and is expressed by equations of the form:

$$10^3 \ln \alpha_b^a = \frac{A \times 10^6}{T^2} + B \quad (17.15)$$

where A and B are constants and T is the temperature in kelvins. Values of A and B for many mineral-water systems were compiled by Friedman and O'Neil (1977) and Faure (1986).

The natural logarithms of small numbers like $1.00x$ have the interesting property that:

$$10^3 \ln 1.00x \approx x \quad (17.16)$$

For example, $10^3 \ln 1.005 = 4.987$; but the approximation deteriorates when $x > 10$, since $10^3 \ln 1.010 = 9.950$. Nevertheless, the approximation does apply to the fractionation of C, N, O, and S in many natural systems, and we can therefore use it to simplify equation 17.15. We start with the equation:

$$\alpha_b^a - 1 = \frac{R_a}{R_b} - 1 = \frac{R_a - R_b}{R_b} \quad (17.17)$$

Next, we use equation 17.11 to relate R_a and R_b to their respective δ -values and substitute them into equation 17.17. After simplifying the resulting expression we obtain:

$$\alpha_b^a - 1 = \frac{R_a - R_b}{R_b} = \frac{(\delta_a + 10^3) - (\delta_b + 10^3)}{(\delta_b + 10^3)} \quad (17.18)$$

which reduces to:

$$\alpha_b^a - 1 = \frac{\delta_a - \delta_b}{\delta_b + 10^3} \quad (17.19)$$

If α_b^a is less than 1.010, $\delta_b + 10^3 \approx 10^3$ and therefore:

$$\alpha_b^a - 1 \approx \frac{\delta_a - \delta_b}{10^3} \quad (17.20)$$

and

$$10^3(\alpha_b^a - 1) \approx \delta_a - \delta_b \quad (17.21)$$

Therefore, we can state that:

$$10^3(\alpha_b^a - 1) \approx 10^3 \ln \alpha_b^a \approx \delta_a - \delta_b \quad (17.22)$$

If we let $\delta_a - \delta_b = \Delta_b^a$, then:

$$10^3 \ln \alpha_b^a \approx \Delta_b^a \approx \frac{A \times 10^6}{T^2} + B \quad (17.23)$$

This equation indicates that the *difference* in the δ -values of C, N, O, and S in coexisting compounds in isotopic equilibrium *decreases* with *increasing* temperature. Ultimately, as T becomes very large, Δ approaches zero because the two coexisting compounds have virtually the *same* isotope composition. However, in the normal temperature range at or near the surface of the Earth, coexisting compounds of H, C, N, O, and S have *different* isotope compositions when the respective exchange reactions are in equilibrium.

17.3 Isotope Fractionation in the Hydrosphere

Evaporation of water from the surface of the oceans in the equatorial region of the Earth causes the formation of air masses containing water vapor that is depleted in ^{18}O and D compared to seawater. In the northern hemisphere the moist air from the tropics is displaced in a northeasterly direction by the circulation of the atmosphere and therefore enters cooler climatic zones. As a result of the decrease in temperature, the air becomes saturated with water vapor, which causes condensation and precipitation of water as rain, snow, or hail.

The condensate is enriched in the heavy isotopes of O and H because condensation is the opposite of evaporation. However, whereas evaporation may take place under disequilibrium conditions, the condensation of water from its vapor occurs at equilibrium. Therefore, we can calculate $\delta^{18}\text{O}$ and δD of the first condensate that forms from water vapor of known isotopic composition at a specified temperature.

According to equation 17.12, the $\delta^{18}\text{O}$ value of the condensate is:

$$\delta^{18}\text{O}_l = \alpha_v^1(\delta^{18}\text{O}_v + 10^3) - 10^3 \quad (17.24)$$

If $\delta^{18}\text{O}_v = -13.1\text{‰}$ relative to SMOW and $\alpha_v^1(\text{O}) = 1.0092$ at 25°C , then:

$$\delta^{18}\text{O}_l = 1.0092(-13.1 + 10^3) - 10^3 = -4.0\text{‰} \quad (17.25)$$

Similarly, the δD value of the condensate that forms at 25°C from vapor having $\delta\text{D}_v = -94.8\text{‰}$ is:

$$\delta\text{D}_l = 1.074(-94.8 + 10^3) - 10^3 = -27.8\text{‰} \quad (17.26)$$

Note that the condensate is enriched in ^{18}O and D relative to the vapor in the air. However, the isotopic composition of the first condensate is not the same as that of the original seawater because we assumed that the isotope fractionation during evaporation was enhanced by kinetic effects.

When the first droplets of condensate fall out of the air mass as rain or snow, the water vapor remaining in the cloud is further depleted in ^{18}O and D. As a result of continuing condensation and loss of water or snow from an air mass, the remaining water vapor becomes progressively more depleted in ^{18}O and D as air masses in the northern hemisphere move in a northeasterly direction. Consequently, the $\delta^{18}\text{O}$ and δD values of *meteoric precipitation* become progressively more negative with increasing geographic latitude.

The resulting *latitude effect* on the isotope composition of meteoric water is caused by (1) progressive isotope fractionation associated with condensation of water vapor and the removal of water droplets from the air mass, (2) the increase of the isotope fractionation factors caused by a decrease in temperature, (3) the reevaporation of meteoric water from the surface of the Earth, and (4) evapotranspiration of water by plants. Since H and O occur together in water molecules and since both experience the same sequence of events during the migration of air masses, the δD and $\delta^{18}\text{O}$ values of meteoric water are strongly correlated and satisfy an empirical equation known as the *meteoric-water line* (Figure 17.1):

$$\delta\text{D} = 8\delta^{18}\text{O} + 10 \quad (17.27)$$

The slope of the meteoric-water line is related to the isotopic fractionation factors of H and O during evaporation or condensation of water:

$$\frac{\alpha_v^1(\text{H}) - 1}{\alpha_v^1(\text{O}) - 1} = \frac{0.074}{0.0092} = 8.0 \quad (17.28)$$

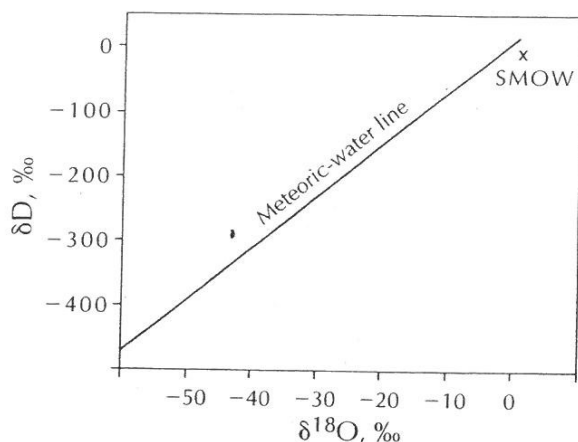


Figure 17.1 Relationship between δD and $\delta^{18}O$ of meteoric precipitation expressed by equation 17.27. Snow in the polar regions is strongly depleted in ^{18}O and D as a result of progressive isotope fractionation of water vapor in air masses moving toward the polar regions.

The intercept at $\delta D = +10\text{‰}$ for $\delta^{18}O = 0\text{‰}$ is called the *deuterium excess*. Both the slope and the intercept vary depending on local climatic conditions (Hoefs, 1996). Nevertheless, meteoric precipitation is progressively depleted in ^{18}O and D with increasing geographic latitude such that winter snow at the South Pole is characterized by $\delta^{18}O = -55\text{‰}$ and $\delta D = -430\text{‰}$.

The isotope fractionation of water vapor in an air mass described above can be modeled by *Rayleigh distillation*, provided that the isotopic composition of the water vapor in the air mass is not modified by evaporation of water from the surface of the Earth, or by evapotranspiration, or by partial evaporation of raindrops falling through the cloud. In accordance with Rayleigh distillation, the isotope ratio (R_v) of the remaining vapor is related to its original isotope ratio (R_v^0) by the equation:

$$R_v = R_v^0 f^{\alpha-1} \quad (17.29)$$

where $\alpha_v^l = R_l/R_v$ and f is the fraction of vapor remaining. We can use equation 17.11 to convert equation 17.29 to the δ -notation:

$$\delta^{18}O_v = (\delta^{18}O_v^0 + 10^3)f^{\alpha-1} - 10^3 \quad (17.30)$$

Similarly, for H:

$$\delta D_v = (\delta D_v^0 + 10^3)f^{\alpha-1} - 10^3 \quad (17.31)$$

When $f = 1.0$ in equation 17.30, $\delta^{18}O_v = \delta^{18}O_v^0$. However, when f varies from 1.0 to 0.0, $\delta^{18}O$ of the *remaining vapor* decreases as shown in Figure 17.2. As the process continues, the isotopic abundances of ^{18}O and D in the *condensate* must also decrease because the isotope ratios of O and H in the liquid are bound to those in the remaining vapor by equation 17.6, that is, $\alpha_v^l = R_l/R_v$. Since $R_v = R_l/\alpha$, we obtain from equation 17.29:

$$R_l = \alpha R_v^0 f^{\alpha-1} \quad (17.32)$$

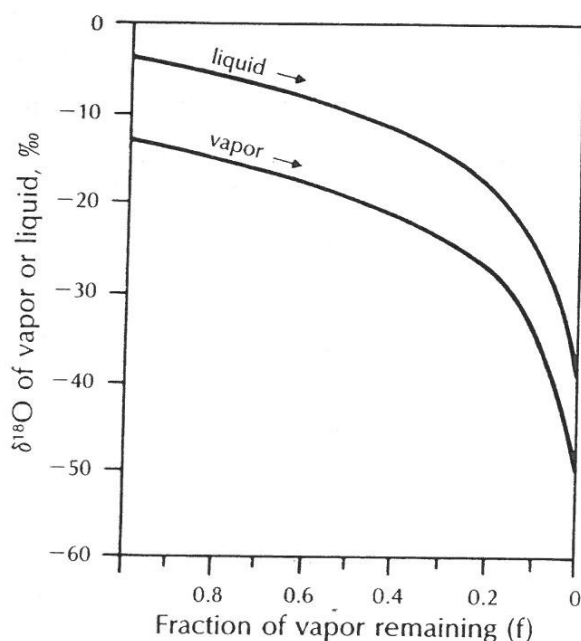


Figure 17.2 Effect of Rayleigh distillation on the $\delta^{18}O$ value of water vapor remaining in an air mass and of meteoric precipitation falling from it at a constant temperature of 25°C and $\alpha_v^l = 1.0092$. Note that meteoric precipitation becomes strongly depleted in ^{18}O as the fraction of vapor remaining in the air mass decreases. Such ^{18}O -depleted snow occurs in the polar regions and at high elevations in mountain ranges, but the effect is probably caused more by condensation at decreasing temperature than by Rayleigh distillation.

and by analogy to equation 17.30:

$$\delta^{18}\text{O}_l = \alpha(\delta^{18}\text{O}_v^0 + 10^3)f^{\alpha-1} - 10^3 \quad (17.33)$$

This equation yields the $\delta^{18}\text{O}$ value of meteoric precipitation in Figure 17.2 that forms by progressive condensation of vapor at constant temperature.

The results help to explain why meteoric water is depleted in ^{18}O and D compared to seawater. However, in reality the process is complicated because of the addition of vapor to the air mass by reevaporation of fresh water and because the isotope fractionation factors of O and H in water are not constant but *increase* with decreasing temperature.

The temperature dependence of the isotope fractionation factor for oxygen during the evaporation/condensation of water from 0 to 100°C has been studied experimentally and is expressed by an equation stated by Friedman and O'Neil (1977):

$$10^3 \ln \alpha = \frac{1.534 \times 10^6}{T^2} - \frac{3.206 \times 10^3}{T} + 2.644 \quad (17.34)$$

where T is the temperature in kelvins. The fractionation of H in water from 0 to 50°C is:

$$10^3 \ln \alpha = \frac{24.844 \times 10^6}{T^2} - \frac{76.248 \times 10^3}{T} + 52.612 \quad (17.35)$$

The condensation of water vapor in air masses to form snow takes place at temperatures below 0°C. The variation of the isotope fractionation factors for ice and its vapor in this temperature range was discussed by Fritz and Fontes (1980b).

The strong dependence of the isotope fractionation factors of O and H in meteoric water on the temperature of evaporation/condensation gives rise to a relationship between the $\delta^{18}\text{O}$ value of meteoric precipitation and average monthly temperatures. The relationship was expressed by Yurtsever (1975) in the form of an empirical equation based on data from northern Europe and Greenland:

$$\delta^{18}\text{O}(\text{SMOW}) = (0.521 \pm 0.014)t - 14.96 \pm 0.21 \quad (17.36)$$

where t is the average monthly temperature in degrees Celsius.

The temperature dependence of $\delta^{18}\text{O}$ and δD values of snow accumulating in the polar regions is reflected in seasonal variations of these parameters, which thereby record long-term changes in climatic conditions. For example, the climatic warming at the Pleistocene-Holocene boundary about 15,000 years ago is recorded in ice cores recovered from Greenland and Antarctica by a 5–10‰ shift in the $\delta^{18}\text{O}$ values of the ice (Johnsen et al., 1972; Moser and Stichler, 1981; Lorius et al., 1985).

17.4 Oxygen Isotope Composition of Calcite

After the end of World War II, Harold C. Urey, the discoverer of deuterium, set up a laboratory at the University of Chicago where he began to study the fractionation of O isotopes between calcite and water. This work soon led to the development of a *paleothermometer* based on the isotope composition of O in marine calcite and aragonite. In addition, Urey assembled a group of gifted students and postdoctoral collaborators whose research has contributed significantly to the systematic study of the isotope compositions of H, C, N, O, and S in terrestrial and extraterrestrial samples. Harold Urey's former colleagues and their students are still at the forefront of the study of isotope fractionation and its applications to geochemistry and cosmochemistry (Valley et al., 1986).

Calcite and aragonite are enriched in ^{18}O compared to the water from which they precipitated. When the reaction takes place in isotopic equilibrium in the oceans, the $\delta^{18}\text{O}$ value of the calcite (δ_c) is related to the $\delta^{18}\text{O}$ of seawater (δ_w) and the temperature (t) in degrees Celsius by the equation:

$$t = 16.9 - 4.2(\delta_c - \delta_w) + 0.13(\delta_c - \delta_w)^2 \quad (17.37)$$

where δ_c is $\delta^{18}\text{O}$ of CO_2 prepared by reacting calcium carbonate with 100% H_3PO_4 at 25°C and δ_w is $\delta^{18}\text{O}$ of CO_2 in equilibrium with the water at 25°C and both δ -values are based on the same standard. Anderson and Arthur (1983) recast equation 17.37 by expressing δ_w on the SMOW scale, whereas δ_c is on the PDB scale:

$$t(^{\circ}\text{C}) = 16.0 - 4.14(\delta_c - \delta_w) + 0.13(\delta_c - \delta_w)^2 \quad (17.38)$$

It is customary to express the $\delta^{18}\text{O}$ values of carbonate minerals relative to the PDB standard (Peedee Formation, belemnite) originally chosen by Urey. Consequently, the $\delta^{18}\text{O}$ values of marine carbonates are close to 0‰ on the PDB scale, whereas those of nonmarine carbonates are negative because fresh water is depleted in ^{18}O .

The $\delta^{18}\text{O}$ value of seawater is included in equations 17.37 and 17.38 because the isotope composition of water in the oceans is variable on a regional scale, depending on the rate of evaporation from the surface, and because of mixing of seawater with fresh water discharged by rivers or by melting icebergs. Moreover, the $\delta^{18}\text{O}$ value of the oceans has varied in the past depending on the extent of continental glaciation. At the height of the Pleistocene glaciation, the oceans became enriched in ^{18}O because the continental ice sheets were strongly depleted in ^{18}O . Therefore, the $\delta^{18}\text{O}$ of seawater has varied by up to 1.5‰ from +0.09‰ on the SMOW scale (full-scale continental glaciation) to ~0‰ (present state) to -0.60‰ (no continental ice sheets).

The oxygen-isotope paleothermometer is also based on the assumption that biogenic calcites or aragonites were precipitated in isotopic equilibrium with seawater. This requirement is met *only* by a few carbonate-secreting organisms including mollusks and the foraminifera. The latter include both benthic (bottom dwelling) and pelagic (floating) species whose habitats have quite different temperatures. Other complications to be considered are the seasonality of shell growth of mollusks and the mineralogy of the shells because calcite and aragonite have slightly different O isotope fractionation factors.

The complications described above make it impractical to use equation 17.37 or 17.38 to determine the numerical values of paleotemperatures in the oceans. Nevertheless, *variations* in the $\delta^{18}\text{O}$ (PDB) values of selected foraminiferal tests from different depths in a sediment core can reveal *changes* of the water temperature in the past at the site where the core was taken. An *increase* in ^{18}O enrichment of calcite may be caused either by a *decrease* in the temperature, or by an *increase* in the ^{18}O content of the water, or both. A *decrease* in the ^{18}O enrichment of calcite may be caused by the opposite changes in the temperature and the $\delta^{18}\text{O}$ value of water. The temperature of the oceans and the $\delta^{18}\text{O}$ value of seawater may be related because a decrease in the temperature of seawater may be caused by the onset of continental glaciation, which also causes the oceans to become enriched in ^{18}O . Unfortunately, equations 17.37 or 17.38 cannot be used to determine both the temperature and the $\delta^{18}\text{O}$ value of seawater because they contain three variables, only one of which can be measured. This difficulty could be overcome by analyzing other authigenic O-bearing minerals such as Ca phosphate or chert. However, the temperature dependence of the isotope fractionation factors for these materials is very similar to that of the carbonates. Therefore, it is not possible to determine both the temperature and isotopic composition of seawater by presently available isotopic methods (Faure, 1986; Hoefs, 1996).

The $\delta^{18}\text{O}$ values of recently formed marine limestones are close to zero on the PDB scale but tend to be slightly negative. The small ^{18}O depletion of calcite in modern coral reef deposits occurs during the recrystallization of the carbonate particles in the presence of fresh water that may accumulate in the subsurface in coral atolls and in coastal reef complexes.

The average $\delta^{18}\text{O}$ values of marine carbonate rocks of Phanerozoic age become increasingly more negative with age (Veizer and Hoefs, 1976). The apparent ^{18}O depletion of these rocks has been attributed to (1) the long-term effects of oxygen-isotope exchange between carbonate rocks

and meteoric water, (2) the existence of *higher* temperatures in the oceans of the Paleozoic Era than at present, and (3) a *lower* ^{18}O content of seawater in Paleozoic time compared to the modern oceans. The explanation for the observed secular changes in the $\delta^{18}\text{O}$ values of marine carbonate rocks may have far-reaching consequences to our understanding of the history of the Earth, but there is as yet no consensus on this important question.

Nonmarine carbonate rocks are variably depleted in ^{18}O compared to marine carbonates and commonly have negative $\delta^{18}\text{O}$ values on the PDB scale because they formed from meteoric water. Some nonmarine carbonate rocks, such as those of the Flagstaff and Green River formations in the western United States, formed by precipitation of calcite in large lakes that became saline as a result of evaporative concentration under arid or semiarid climatic conditions. If the precipitates formed in isotopic equilibrium, the fractionation of O isotopes between calcite and water varies with temperature between 0 and 500°C in accordance with the equation (Friedman and O'Neil, 1977):

$$10^3 \ln \alpha_w^c = \frac{2.78 \times 10^6}{T^2} - 2.89 \quad (17.39)$$

where letters c and w refer to calcite and water, respectively, and T is the temperature in kelvins. This equation can be used to determine the $\delta^{18}\text{O}$ value of the water from which calcite precipitated, as illustrated in the following example.

Calcite concretions in the Pagoda Tillite of Permian age in the Queen Alexandra Range of the Transantarctic Mountains have an average $\delta^{18}\text{O}$ value of about -20‰ on the PDB scale (Lord et al., 1988). In order to calculate $\delta^{18}\text{O}$ of the water from which the concretion was deposited, we convert $\delta^{18}\text{O}$ of the concretion from the PDB to the SMOW scale by means of equation (17.10).

$$\begin{aligned} \delta^{18}\text{O}_c (\text{SMOW}) &= 1.03086(-20) + 30.86 \\ &= +10.2\text{‰} \end{aligned} \quad (17.40)$$

The temperature at which the calcite precipitated was probably just above the freezing point

because tillite is a glacial deposit. If the temperature was 0°C , equation 17.39 yields the fractionation factor for oxygen in calcite and water:

$$10^3 \ln \alpha_w^c (\text{O}) = \frac{2.78 \times 10^6}{(273.15)^2} - 2.89 = 34.36 \quad (17.41)$$

Therefore, $\ln \alpha_w^c (\text{O}) = 0.03436$ and $\alpha_w^c (\text{O}) = 1.03495$. Next, we use equation 17.12 to calculate $\delta^{18}\text{O}_w$:

$$1.03495 = \frac{10.2 + 10^3}{\delta^{18}\text{O}_w + 10^3} \quad (17.42)$$

which yields $\delta^{18}\text{O}_w = -23.9\text{‰}$ on the SMOW scale. The approximation represented by equation 17.22:

$$\delta_c - \delta_w \approx 10^3 \ln \alpha_w^c \quad (17.43)$$

leads to $\delta_w = 10.2 - 34.36 = -24.16\text{‰}$, which differs from the precise value (-23.9‰) by only about 1%.

The results of this calculation indicate that the concretions in the Pagoda Tillite of Antarctica were deposited by water that was strongly depleted in ^{18}O . This result implies that the water formed by melting of the Gondwana ice sheet, which also deposited the sediment of the Pagoda Tillite. Evidently, the $\delta^{18}\text{O}$ value of the Gondwana ice sheet, which existed more than 245 million years ago, has been preserved in the calcite concretion. The isotopic composition of H in the glacial meltwater is given by equation 17.27:

$$\delta\text{D} = 8(-23.9) + 10 = -181.2\text{‰} \quad (17.44)$$

The O and H isotope compositions of the Gondwana ice sheet are similar to those of snow that is presently forming in the coastal areas of Antarctica and in Greenland. The average monthly temperature of the area of accumulation of the Gondwana ice sheet can be estimated from equation 17.36:

$$-23.9 = 0.521t - 14.96 \quad (17.45)$$

which yields $t = -17.2^\circ\text{C}$. This temperature applies to the area of *accumulation* of the Gondwana ice sheet and not to its margin in Antarctica where the ice was melting during the Permian Period.

17.5 Oxygen and Hydrogen in Clay Minerals

Oxygen in the silicate and aluminosilicate minerals of igneous and metamorphic rocks and their sedimentary derivatives is enriched in ^{18}O relative to SMOW. When these aluminosilicate minerals react with acidified water to form kaolinite or other clay minerals (Chapter 16), the isotopic compositions of O and H in the clay minerals depend on several factors:

1. The isotope composition of the water.
2. The isotopic composition of the reactant minerals.
3. The amount of water compared to the amount of reactant minerals (water/rock ratio).
4. The temperature.
5. The isotopic fractionation factors of O and H.

The reactions between feldspar and acidified meteoric water to form clay minerals take place in the presence of an excess of water such that

the water/rock ratio is large. Therefore, the isotope composition of O and H in clay minerals depends on the water rather than on the reactant minerals. Since $\delta^{18}\text{O}$ and δD of meteoric water vary depending on geographic latitude and climatic temperatures, the clay minerals formed on the continents inherit the relationship between δD and $\delta^{18}\text{O}$ expressed by the meteoric-water line. However, clay minerals are enriched in ^{18}O and depleted in D relative to the water at the site of the reaction. Therefore, the equations relating $\delta^{18}\text{O}$ and δD of clay minerals have similar slopes but different intercept values than the meteoric-water line. According to Savin and Epstein (1970), the equations are:

$$\delta\text{D} = 7.3\delta^{18}\text{O} - 260 \quad (\text{montmorillonite}) \quad (17.46)$$

$$\delta\text{D} = 7.5\delta^{18}\text{O} - 220 \quad (\text{kaolinite}) \quad (17.47)$$

These equations have been plotted in Figure 17.3 in order to illustrate the relationship between the isotopic composition of clay minerals and meteoric water.

When clay minerals that formed on land are transported and redeposited in the oceans, the

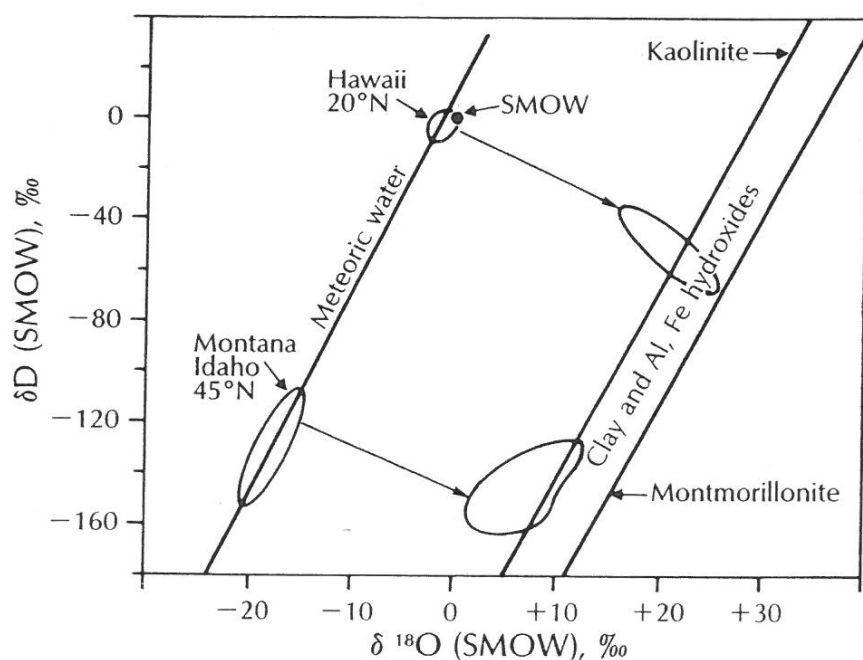


Figure 17.3 Schematic representation of the isotope compositions of O and H in clay minerals and Al, Fe hydroxides of soils and the local meteoric water. The meteoric-water line is equation 17.27, whereas the montmorillonite and kaolinite lines are equations 17.46 and 17.47, respectively. The isotope compositions of meteoric water and soil from Hawaii and the Montana-Idaho area of the United States illustrate the ^{18}O enrichments and D depletion of clay minerals relative to local meteoric precipitation. The diagram also depicts the latitude effect on the isotopic composition of meteoric precipitation (based on data by Lawrence and Taylor, 1971).

isotope compositions of H and O remain unchanged. Other detrital minerals associated with clay minerals in marine muds likewise do not equilibrate isotopically at the time of deposition. The oxygen isotope fractionation factors of clay minerals at Earth-surface temperatures listed in Table 17.2 range from 1.0236 for illite to 1.027 for kaolinite and montmorillonite. Therefore, the $\delta^{18}\text{O}$ value of kaolinite that forms from microcline in the presence of meteoric water having $\delta^{18}\text{O} = -10\text{‰}$ is given by equation 17.12:

$$\delta^{18}\text{O}(\text{kaolinite}) = 1.027(-10 + 10^3) - 10^3 \\ = +16.7\text{‰} \quad (17.48)$$

The corresponding δD value is obtained from equation 17.46:

$$\delta\text{D} = 7.3(+16.7) - 260 = -138.1\text{‰} \quad (17.49)$$

These values are typical of clay-rich soils as indicated in Figure 17.3 based on measurements by Lawrence and Taylor (1971).

The mineral particles of *clastic* sedimentary rocks approach isotopic equilibrium *only* at elevated temperatures when the rates of the respective isotope exchange reactions become significant. According to O'Neil and Kharaka (1976), temperatures of $\sim 100^\circ\text{C}$ are required to equilibrate H, whereas a temperature of 300°C or more is required to equilibrate O between clay minerals and water. However, note that the achievement of isotopic equilibrium by a clastic sedimentary rock

does *not* mean that all minerals then contain O or H of the same isotope composition.

In order to illustrate this point we use the data in Table 17.3 to calculate O-isotope fractionation factors and $\delta^{18}\text{O}$ values of detrital illite, smectite, quartz, and calcite in contact with seawater ($\delta^{18}\text{O} = 0.0\text{‰}$) at 100°C , assuming that sufficient time is available to achieve isotope equilibrium.

Mineral	α_1^c	$\delta^{18}\text{O}, \text{‰}$
Illite	1.0127	+12.7
Smectite	1.0144	+14.4
Quartz	1.0210	+21.1
Calcite	1.0167	+16.7

Evidently, each mineral has a different $\delta^{18}\text{O}$ value even though all were assumed to have equilibrated O with water of the same isotopic composition at the same temperature.

Clastic sedimentary rocks are generally composed of mineral particles whose H and O isotope compositions may have been only partly reset

Table 17.2 Isotope Fractionation Factors for Hydrogen and Oxygen Between Clay Minerals (Including Gibbsite) and Water at Earth-Surface Temperatures

Mineral	Hydrogen	Oxygen
Montmorillonite	0.94	1.027
Kaolinite	0.97	1.027
Glauconite	0.93	1.026
Illite	—	1.0236
Gibbsite	0.984	1.018

SOURCE: Faure (1986), Table 25.3.

Table 17.3 Temperature Dependence of the O-Isotope Fractionation Factor Between Minerals and Water in the Form $10^3 \ln \alpha = A \times 10^6/T^2 + B^a$

Mineral	A	B
Clay Minerals		
Kaolinite	2.5	-2.87
Illite	2.43	-4.82
Smectite	2.67	-4.82
Chlorite	1.56	-4.70
Other Minerals		
Quartz	3.38	-3.40
Calcite	2.78	-2.89 ^b
Dolomite	3.14	-2.0
Anhydrite	3.21	-4.72

^aCompiled from the literature and recalculated by Anderson and Arthur (1983) to be consistent with Friedman and O'Neil (1977).

^bFrom Friedman and O'Neil (1977).

during burial and diagenesis depending on the temperature and available time. The $\delta^{18}\text{O}$ values of marine shales range from +14 to +19‰ (Savin and Epstein, 1970), whereas their δD values lie between -33 and -73‰ (Yeh, 1980). The $\delta^{18}\text{O}$ values of marine limestones decrease with their depositional age from about +30‰ (SMOW) for Quaternary rocks to +21‰ (SMOW) in specimens of Cambrian age (Keith and Weber, 1964; Veizer and Hoefs, 1976). Since Ca carbonate does not contain H, the δD values of limestones are not defined.

Note that this section deals with systems having *large* water/rock ratios. Under such conditions the water contains a virtually infinite reservoir of H and O whose isotope compositions are not affected by exchange reactions with minerals in contact with them. Such environments occur primarily on the surface of the Earth where mineral particles are exposed to large quantities of water.

The situation *below* the surface of the Earth is commonly quite different because in this environment relatively small quantities of water are in contact with large volumes of rock for long periods of time and at elevated temperatures. In such cases the water/rock ratio is *small* and therefore the isotope compositions of O and H in the water change significantly, whereas those of the rocks change much less.

17.6 Groundwater and Geothermal Brines

The isotope compositions of O and H in groundwater are altered by isotope exchange reactions with the rocks in which the water occurs. The magnitude of this effect increases primarily with temperature and with the subsurface residence age of the water, but depends also on the mineral composition of the rocks and on the presence of other substances such as hydrocarbons and $\text{H}_2\text{S}(\text{g})$. Several of these environmental properties, including temperature and residence age of the water, may increase with depth below the surface.

Water occurring in the pore spaces and fractures of rocks in the subsurface may originate

from several different sources defined by Kharaka and Carothers (1986):

Meteoric water was derived from rain, snow, water courses, and other bodies of surface water that percolates through pore spaces in the rocks and displaces the interstitial water they contain regardless of its origin.

Connate water was deposited with sediments and has been out of contact with the atmosphere since the time of deposition. In the case of sedimentary rocks of marine origin the connate water was derived from seawater trapped in the pore spaces of the sediment.

Diagenetic water is released from solid phases as a result of reactions among minerals during diagenesis, such as the transformation of gypsum to anhydrite and smectite to illite both of which release water.

Formation water is the water that exists in layers of sedimentary rocks prior to drilling regardless of its origin.

Brine is water whose salinity, expressed in terms of total dissolved solids (TDS), exceeds that of seawater, which is 35,000 mg/L.

When meteoric water moves through the pore spaces of subsurface rocks, isotope exchange reactions take place that alter the isotope composition of this water in a systematic way. Figure 17.4 illustrates the case where meteoric water having $\delta^{18}\text{O} = -10$ ‰, $\delta\text{D} = -70$ ‰ on the SMOW scale interacts with shale having $\delta^{18}\text{O} = +14$ ‰ and δD between -33 and -73‰ (SMOW). If the water/rock ratio is small, the $\delta^{18}\text{O}$ and δD values of the water change along linear trajectories toward the isotope composition of the shale. The location of a particular water sample on such a trajectory depends primarily on the temperature, which controls the values of the isotope fractionation factors. Since the fractionation factors *decrease* with increasing temperature, $\delta^{18}\text{O}$ and δD of the water move closer *toward* the isotope composition of the rocks as the equilibrium temperature *rises*.

In order to illustrate this relationship we consider two points, A and B, in Figure 17.4, which

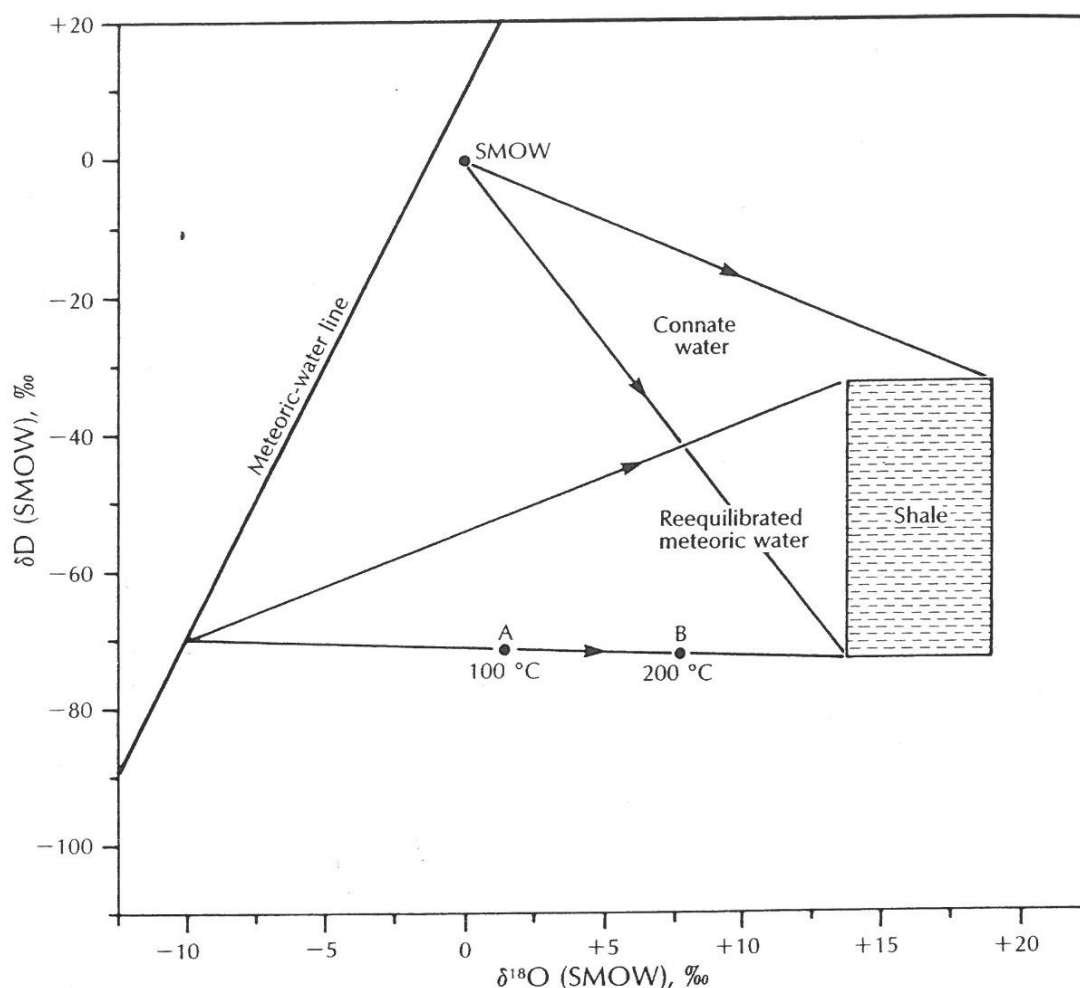


Figure 17.4 Schematic view of the changes in the isotopic composition of groundwater caused by isotope reequilibration with rocks at increasing temperature in systems having low water/rock ratios. Points *A* and *B* are samples of meteoric water ($\delta^{18}\text{O} = -10\text{‰}$) that reequilibrated with illite at 100°C (point *A*) and 200°C (point *B*) assuming that $\delta^{18}\text{O}$ of illite is $+14\text{‰}$ (SMOW) and remains constant. The diagram shows that both meteoric water and seawater may be altered isotopically by this process. As a result of isotope reequilibration groundwater may be displaced from the meteoric-water line to form linear trajectories directed toward the isotope composition of the host rocks. Alternatively, such trajectories may also be the result of mixing between old isotopically reequilibrated brines and meteoric water from the surface of the Earth.

represent meteoric water that has equilibrated O with illite at 100 and 200°C , respectively. The $\delta^{18}\text{O}$ of the water was calculated from the appropriate equation in Table 17.3:

$$10^3 \ln \alpha_w^i = \frac{2.43 \times 10^6}{T^2} - 4.82 \quad (17.50)$$

where α_w^i is the fractionation factor for O between illite (*i*) and water (*w*). We assume that the water/rock ratio is small, that isotopic equilibrium between illite and water has been established, and that $\delta^{18}\text{O}$ of illite is $+14\text{‰}$ (SMOW). The $\delta^{18}\text{O}$ of the water can be calculated by solving equation 17.50 for α_w^i at the desired temper-

atures and by substituting these values into equation 17.12. The results are $\delta^{18}\text{O}_w = +1.3\text{‰}$ at 100°C (point A in Figure 17.4) and $\delta^{18}\text{O}_w = +7.9\text{‰}$ at 200°C (point B). Evidently, $\delta^{18}\text{O}$ of the meteoric water is shifted from -10‰ at the surface toward $\delta^{18}\text{O}$ of the rocks by an amount that increases with increasing temperature.

In the case illustrated in Figure 17.4 the isotope composition of H of the meteoric water changes very little because the water and the rock were assumed to have similar δD values. In general, the δD values of reequilibrated groundwaters do change with increasing temperature depending on the presence of H-bearing compounds, including not only clay minerals but also hydrocarbons and H_2S . However, the temperature dependence of the applicable fractionation factors for H is not yet known.

We conclude that the isotopic compositions of groundwater derived from meteoric precipitation may be displaced from the meteoric-water line by enrichment in ^{18}O and D. Water samples collected at increasing depth below the surface tend to lie along linear trajectories directed toward the $\delta^{18}\text{O}$ and δD values of the subsurface rocks with which the water is equilibrating at varying temperatures depending on the depth. Such linear trajectories are commonly observed by isotope analyses of oilfield brines (Clayton et al., 1966; Hitchon and Friedman, 1969).

Seawater originally trapped in the pore spaces of sediment is likewise altered isotopically by isotope exchange reactions with the host rocks. Figure 17.4 outlines the area in the $\delta^{18}\text{O}$ – δD plane where such waters are located, provided they equilibrated with shale over a range of elevated temperatures in environments having low water/rocks ratios. Therefore, connate water in rock formations of marine origin is ancient seawater whose isotopic and chemical composition has changed as a result of water–rock interactions. Formation waters may be mixtures of connate water and meteoric water, both of which had previously reequilibrated H and O depending on the composition of the rocks, the temperature, and residence age of the water.

Evidently, the isotope composition of groundwater can be understood also as a result of mixing of waters of differing compositions and origins. Meteoric water percolating downward from the Earth's surface may mix at depth with brines that may themselves be mixtures of connate water and highly reequilibrated old meteoric water. The resulting water compositions also lie in the area of reequilibrated meteoric water in Figure 17.4. Therefore, mixing is an important process affecting the isotopic and chemical composition of groundwater in ways that may be indistinguishable from the results of progressive chemical and isotopic equilibration between rocks and water in the subsurface (Lowry et al., 1988).

17.7 Isotope Fractionation of Carbon

Carbon is the principal element in the biosphere, but it occurs also in the lithosphere, hydrosphere, and atmosphere of the Earth. It has two stable isotopes (^{12}C and ^{13}C , Table 17.1) whose abundances in terrestrial samples vary by $\sim 10\%$ as a result of isotope fractionation by exchange reactions and kinetic effects. The isotopic composition of carbon is expressed in terms of the $\delta^{13}\text{C}$ parameter defined by equation 17.9 relative to the PDB standard.

The fractionation of C isotopes occurs primarily during *photosynthesis* reactions in green plants. As a result, biogenic C compounds are strongly *depleted* in ^{13}C and have negative $\delta^{13}\text{C}$ values. However, CaCO_3 precipitated in equilibrium with the CO_2 of the atmosphere is *enriched* in ^{13}C relative to CO_2 and has $\delta^{13}\text{C}$ values near 0‰ because the PDB standard is itself a marine carbonate. However, under certain circumstances, biogenic C may be oxidized to bicarbonate ions which may be incorporated into CaCO_3 , giving it negative $\delta^{13}\text{C}$ values whose magnitude depends on the proportions of mixing of biogenic and inorganic C.

The photosynthesis process starts with diffusion of CO_2 gas from the atmosphere into the cells of green plants and causes the internal CO_2 to be *depleted* in ^{13}C . The internal CO_2 is subse-

quently converted into phosphoglyceric acid containing three C atoms (Calvin cycle) or into dicarboxylic acid with four C atoms (Hatch-Slack cycle). These transformations are irreversible and cause further depletion in ^{13}C amounting to 17–40‰ or more in C_3 plants but only 2–3‰ in C_4 plants. The subsequent synthesis of organic compounds causes additional isotope fractionation with maximum ^{13}C -depletion in lipids (oils, fats, and waxes). Consequently, those land plants which use the C_3 metabolic process are strongly depleted in ^{13}C and have $\delta^{13}\text{C}$ values between –23 and –34‰ (PDB), whereas C_4 plants (corn and tropical grasses) are less depleted in ^{13}C than C_3 plants, and their $\delta^{13}\text{C}$ values range from –6 to –23‰ (PDB). Note that the $\delta^{13}\text{C}$ values of C_3 and C_4 plants do not overlap (Deines, 1980).

Fossil fuels (coal, petroleum, and natural gas) are strongly depleted in ^{13}C because they originated primarily from the remains of C_3 plants. The $\delta^{13}\text{C}$ values of coal range from about –20 to –30‰ (PDB) and average about –25‰. Petroleum is slightly more depleted in ^{13}C than coal, and its $\delta^{13}\text{C}$ values cluster between –27 and –30‰ (PDB).

The isotopic composition of hydrocarbon gases depends on their molecular composition and origin. In general, methane (CH_4) is the most abundant constituent of natural gas and tends to be more depleted in ^{13}C than ethane (C_2H_6), propane (C_3H_8), or butane (C_4H_{10}). Methane formed by bacterial decomposition of cellulose under reducing conditions in soil and recently deposited sediment has strongly negative $\delta^{13}\text{C}$ values ranging to about –80‰ (PDB). Methane produced by cracking of larger hydrocarbon molecules or during the formation of coal or petroleum is depleted in ^{13}C relative to its source material, such that its $\delta^{13}\text{C}$ values vary between –25 and –50‰ (PDB). Deines (1980) showed that methane recovered from relatively shallow depths between 100 and ~2000 m is more depleted in ^{13}C than methane that originated from greater depths. The greater ^{13}C -depletion of methane originating from near-surface sources may be caused by the presence of varying proportions of bacteriogenic methane.

Carbon dioxide of the atmosphere has a $\delta^{13}\text{C}$ value of –7.0‰, but its isotopic composition is

changing because of the addition of biogenic C released by the combustion of fossil fuel. When CO_2 dissolves in water to form carbonic acid and its ions, isotope exchange reactions occur that enrich HCO_3^- in ^{13}C relative to atmospheric CO_2 (Table 17.4). When CaCO_3 precipitates from such solutions the solid phase is further enriched in ^{13}C . According to Deines et al. (1974), the resulting fractionation of C-isotopes in CaCO_3 relative to CO_2 varies with the temperature such that:

$$10^3 \ln \alpha_g^s = \frac{1.194 \times 10^6}{T^2} - 3.63 \quad (17.51)$$

where s and g refer to solid CaCO_3 and gaseous CO_2 , respectively. Therefore, the isotope fractionation factor for C in the CaCO_3 – CO_2 system at 20°C is 1.0103. If $\delta^{13}\text{C}$ of the CO_2 = –7.0‰, then CaCO_3 formed in isotopic equilibrium at 20°C has $\delta^{13}\text{C}$ = +3.2‰ (PDB) based on equation 17.12.

The $\delta^{13}\text{C}$ values of *marine* carbonate rocks of Phanerozoic age are actually close to 0‰. For example, Keith and Weber (1964) reported an average value of $+0.56 \pm 1.55$ ‰ (PDB) for 321 selected samples of marine limestones of Phanerozoic age. The C in *nonmarine* carbonate rocks tends to be depleted in ^{13}C and has more variable isotope compositions than marine carbonates. Keith and Weber (1964) reported an average $\delta^{13}\text{C}$ = -4.93 ± 2.75 ‰ (PDB) for 183

Table 17.4 Fractionation of Carbon Isotopes Between Aqueous Carbonate Ions and Calcium Carbonate Relative to $\text{CO}_2(\text{g})$

Carbonate species	A^a	B^a
H_2CO_3^0	0.0063	–0.91
HCO_3^-	1.099	–4.54
CO_3^{2-}	0.87 (approx.)	–3.4 (approx.)
$\text{CaCO}_3(\text{s})$	1.194	–3.63

^a A and B are constants in the equation

$$10^3 \ln \alpha = \frac{A \times 10^6}{T^2} + B$$

where α is defined as $R(\text{carbonate species})/R(\text{CO}_2(\text{g}))$ and R is the atomic $^{13}\text{C}/^{12}\text{C}$ ratio.

SOURCE: Deines et al. (1974).

selected samples of nonmarine limestones. The ^{13}C -depletion of nonmarine carbonate rocks is attributable to the presence of biogenic carbon formed by oxidation of plant debris to HCO_3^- in the environment of deposition. However, nonmarine carbonates deposited where no plant debris was present have $\delta^{13}\text{C}$ values that are indistinguishable from those of marine carbonates (Lord et al., 1988).

Calcite cements in sandstones may also be depleted in ^{13}C where CH_4 escapes from petroleum pools or coal seams and is oxidized to HCO_3^- (Donovan et al., 1974). Similarly, the calcite in the cap rocks of salt domes has negative $\delta^{13}\text{C}$ values because it precipitated from HCO_3^- ions that formed by the bacterial oxidation of petroleum in pools along the flanks of the salt domes. Cheney and Jensen (1967) reported an average $\delta^{13}\text{C} = -36.2 \pm 6.2\%$ (PDB) for calcites associated with salt domes in Louisiana and Texas. They also found ^{13}C -depleted calcites in the sandstone-hosted U deposits in the Wind River Formation (Lower Eocene), Gas Hills, Wyoming.

Isotope analyses of O and C in calcite can now be made by releasing CO_2 from "zap pits" or microscopic traverses using a laser. Hence, it is possible to study mineral cements or carbonate concretions in great detail on a scale of about $10\text{ }\mu\text{m}$ by means of laser-released CO_2 gas.

17.8 Isotope Compositions of Strontium in Carbonate Rocks

Carbonate rocks contain not only O and C but also include trace amounts of Sr whose isotope composition is variable because of the decay of naturally occurring radioactive ^{87}Rb to stable ^{87}Sr (Section 16.3b). Therefore, we digress briefly to discuss the isotope composition of Sr in carbonate rocks even though the isotopes of Sr are *not* fractionated by exchange reactions or by kinetic effects (Faure, 1986).

Calcium carbonate permits Sr^{2+} to replace Ca^{2+} but excludes Rb^+ . As a result, calcite has a very low Rb/Sr ratio, and its $^{87}\text{Sr}/^{86}\text{Sr}$ ratio is

not significantly altered by radioactive decay of ^{87}Rb to ^{87}Sr after deposition (Section 16.3b). Instead, the isotope composition of Sr in calcite deposited in the oceans results from mixing of different isotopic varieties of Sr that enter the oceans. The principal sources of marine Sr having distinctive $^{87}\text{Sr}/^{86}\text{Sr}$ ratios are (1) old granitic basement rocks of the continental crust (high Rb/Sr, high $^{87}\text{Sr}/^{86}\text{Sr}$); (2) young volcanic rocks along midocean ridges, in oceanic islands, and along continental margins (low Rb/Sr, low $^{87}\text{Sr}/^{86}\text{Sr}$); and (3) marine carbonate rocks on the continents (low Rb/Sr, intermediate $^{87}\text{Sr}/^{86}\text{Sr}$).

The isotopes of Sr are not fractionated during precipitation of calcite or aragonite from aqueous solutions because the mass difference between ^{87}Sr and ^{86}Sr (defined in Table 17.1) is only 1.2%. Because of the absence of isotope fractionation effects and the negligibly small production of radiogenic ^{87}Sr by decay of ^{87}Rb in carbonate rocks, we conclude that marine and nonmarine carbonate rocks, as well as calcite cements and concretions, record the isotope composition of Sr in the *fluid phase* at the time of deposition.

Isotope analyses of Sr in a large number of marine limestones of Precambrian and Phanerozoic age have revealed that the $^{87}\text{Sr}/^{86}\text{Sr}$ ratio of seawater has varied systematically with time (Burke et al., 1982). These variations must have been caused by changes in the isotopic compositions of Sr that entered the oceans from various sources and by changes in the relative proportions of these inputs. The systematic variation of the isotope composition of Sr in the oceans therefore reflects the operation of the "geochemical machine" mentioned first in Section 14.5. However, we still do not have a complete explanation of the specific reasons for the way in which the $^{87}\text{Sr}/^{86}\text{Sr}$ ratio in the oceans has varied with time (Faure, 1986).

The steady rise in the $^{87}\text{Sr}/^{86}\text{Sr}$ ratios of the oceans during the Tertiary period is shown in Figure 17.5 based on measurements by DePaolo and Ingram (1985). This phenomenon can be used for dating marine carbonate samples of Tertiary age and for stratigraphic correlations (Hurst, 1986). The isotope composition of Sr in the modern oceans is constant throughout the oceans

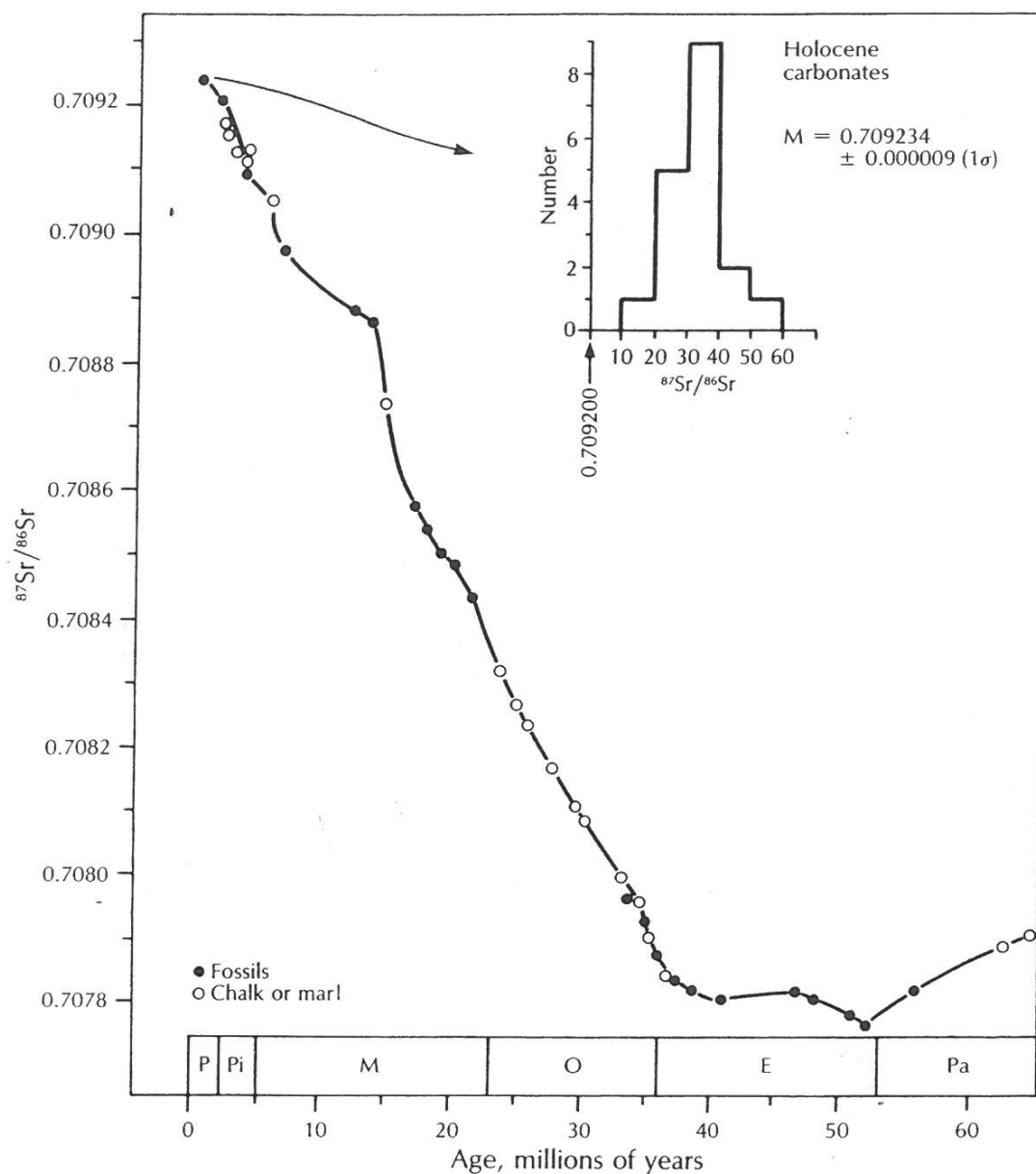


Figure 17.5 Variation of the $^{87}\text{Sr}/^{86}\text{Sr}$ ratio of marine carbonates with age during the Tertiary Period based on measurements by DePaolo and Ingram (1986). The continuous increase of this ratio since the end of the Eocene epoch can be used to date samples of marine carbonate of Tertiary age. Note that the rate of change of the $^{87}\text{Sr}/^{86}\text{Sr}$ ratio has varied with time and that the present value of this ratio in the oceans is 0.709234 ± 0.000009 . However, the $^{87}\text{Sr}/^{86}\text{Sr}$ ratio measured by DePaolo and Ingram (1986) are *higher* than those reported by Burke et al. (1982) by 0.00017. The time scale includes P = Pleistocene, Pi = Pliocene, M = Miocene, O = Oligocene, E = Eocene, and Pa = Paleocene.

because its residence time is ~ 4000 times longer than the time required for mixing. Measurements by DePaolo and Ingram (1986) indicate that the average $^{87}\text{Sr}/^{86}\text{Sr}$ ratio of modern seawater is $0.709\,234 \pm 0.000\,009$. However, this value must be reduced by 0.000 17 to make it compatible with measurements by Burke et al. (1982) and with a compilation of data from the literature by Faure (1982).

The $^{87}\text{Sr}/^{86}\text{Sr}$ ratios of *nonmarine* limestones reflect the ages and Rb/Sr ratios of the rocks exposed to weathering in the drainage basin in which the carbonates were deposited. Therefore, stratigraphic variations of the $^{87}\text{Sr}/^{86}\text{Sr}$ ratios of nonmarine carbonate rocks can only be caused by *changes* in the geology or hydrology of the drainage basin (Neat et al., 1979).

The $^{87}\text{Sr}/^{86}\text{Sr}$ ratios of carbonate cements in arkosic or lithoclastic sandstones and conglomerates are affected by the release of ^{87}Sr from Rb-bearing minerals in the sediment during diagenesis. Therefore, the homogeneity of $^{87}\text{Sr}/^{86}\text{Sr}$ ratios of calcite cements in a lithologic unit is a reflection of the rate of flow of formation water through the lithologic unit or aquifer. Relatively rapid throughput of water tends to make $^{87}\text{Sr}/^{86}\text{Sr}$ ratios homogeneous, whereas stagnant conditions caused by low permeability or a low hydrostatic gradient permit local variations to develop in the isotope composition of Sr of calcite or zeolite cements (Stanley and Faure, 1979).

17.9 Isotope Fractionation of Sulfur

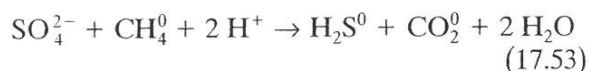
The stable isotopes of S listed in Table 17.1 are fractionated during the reduction of sulfate ions by bacteria and as a result of isotope exchange reactions among S-bearing ions and molecules in aqueous solution. The isotope composition is expressed in the δ -notation as:

$$\delta^{34}\text{S} = \left(\frac{(^{34}\text{S}/^{32}\text{S})_{\text{spl}} - (^{34}\text{S}/^{32}\text{S})_{\text{std}}}{(^{34}\text{S}/^{32}\text{S})_{\text{std}}} \right) \times 10^3 \quad (17.52)$$

where the standard is the S in troilite (FeS) of the iron meteorite Canyon Diablo collected at

Meteor Crater, Arizona. The isotope composition of S is determined by analyzing SO_2 gas prepared from samples on a multicollector gas-source mass spectrometer (Hoefs, 1996).

The most important process causing isotope fractionation of S is associated with the metabolism of the bacteria *Desulfovibrio* and *Desulfatamaculum*. These bacteria flourish in anoxic environments by oxidizing organic matter using O derived from sulfate ions. In this process, the S is reduced from +6 to -2 and is subsequently expelled as H_2S , which is enriched in ^{32}S compared to the sulfate. We can approximate this transformation by constructing an equation to represent this oxidation-reduction reaction occurring without the aid of bacteria:



The standard free energy change of this reaction is $\Delta G_R^\circ = -26.324$ kcal and its equilibrium constant at 25°C is $K = 10^{19.30}$. Evidently, SO_4^{2-} and aqueous CH_4 react to form H_2S^0 , CO_2^0 , and H_2O , which dominate at equilibrium. However, the reduction of SO_4^{2-} by bacteria is not at equilibrium and the isotope fractionation of S is caused by kinetic effects (Fry et al., 1986).

The extent of fractionation of S isotopes by bacteria is inversely proportional to the rate of metabolism and depends on the composition and abundance of the food supply, on the size of the sulfate reservoir, on the temperature, and on the rate of removal of the H_2S . In general, bacterial isotope fractionation of S depends on the environmental conditions represented by three alternative scenarios:

1. In *closed* systems, which become anoxic because vertical mixing is inhibited, the concentration of H_2S may rise until it ultimately poisons the bacteria. Such environments exist in the Black Sea, in isolated depressions in the oceans, and in some lakes where H_2S is strongly depleted in ^{34}S .
2. In open systems in which the sulfate reservoir is virtually infinite, but H_2S is continually removed by degassing or by precipitation

as insoluble sulfides, the metabolic S is depleted in ^{34}S but the isotope composition of S remains nearly constant.

3. Systems containing a limited amount of sulfate, but from which the metabolic H_2S can escape, are subject to Rayleigh distillation. Therefore, both the remaining sulfate as well as the H_2S become enriched in ^{34}S because of the systematic and irreversible loss of ^{32}S from the system.

The metabolic H_2S excreted by bacteria in anoxic marine and nonmarine basins may be precipitated as hydrotroilite ($\text{FeS} \cdot n\text{H}_2\text{O}$), which ultimately recrystallizes as pyrite (FeS_2). The S in such sedimentary sulfide minerals is generally depleted in ^{34}S , but the $\delta^{34}\text{S}$ values may vary stratigraphically, and may even become positive, depending on the extent of Rayleigh distillation. The wide range of variation of $\delta^{34}\text{S}$ values and the commonly observed depletion in ^{34}S distinguish sedimentary sulfides from those in igneous rocks whose $\delta^{34}\text{S}$ values cluster close to 0.0‰. However, the isotope composition of S alone is *not* a reliable criterion for distinguishing between sulfides of sedimentary and magmatic origin because:

1. The $\delta^{34}\text{S}$ values of some sedimentary sulfides may be near 0.0‰ or may even be positive.
2. The $\delta^{34}\text{S}$ values of some magmatic sulfides may be negative because the S originated from sedimentary rocks in the vicinity of igneous intrusions.
3. The $\delta^{34}\text{S}$ values of sulfide minerals in metamorphosed igneous or sedimentary rocks may be altered by isotope exchange reactions at elevated temperatures.

When sulfide minerals form by precipitation from a hot aqueous fluid or by crystallization from immiscible sulfide liquids in magmas, the S they contain may have different isotope compositions because of isotope exchange reactions with a common S reservoir. The magnitude of the *differences* in the $\delta^{34}\text{S}$ values of pairs of coexisting sulfide minerals *increases* with decreasing temperature. If the relevant isotope exchange reac-

tions achieved equilibrium, differences in $\delta^{34}\text{S}$ values of coexisting sulfide minerals may indicate the isotope equilibration temperature based on the experimentally determined calibration equations listed in Table 17.5.

Fractionation of S isotopes between sphalerite and galena has been investigated several times with somewhat different results. For example, at 200°C the equation of Grootenboer and Schwarcz (1969) yields $\alpha_{\text{gn}}^{\text{sp}} = 1.00285$, whereas that of Kajiura and Krouse (1971) yields $\alpha_{\text{gn}}^{\text{sp}} = 1.00360$, where sp and gn refer to sphalerite and galena, respectively. Since in the example cited above $(\alpha - 1)10^3 < 10$, we can use the approximation:

$$10^3 \ln \alpha_{\text{gn}}^{\text{sp}} = \delta^{34}\text{S}_{\text{sp}} - \delta^{34}\text{S}_{\text{gn}} = \Delta_{\text{gn}}^{\text{sp}} \quad (17.54)$$

Ohmoto and Rye (1979) used this relationship to express the temperature dependence of S-isotope fractionation between sphalerite and galena:

$$\Delta_{\text{gn}}^{\text{sp}} = (7.2 \times 10^5)/T^2 \quad (17.55)$$

Evidently, $\Delta_{\text{gn}}^{\text{sp}}$, which is defined as $\delta^{34}\text{S}(\text{sphalerite}) - \delta^{34}\text{S}(\text{galena})$, *decreases* with *increasing* equilibration temperature. When $\Delta_{\text{gn}}^{\text{sp}} = 4.0$, equation 17.55 yields:

$$T = \left(\frac{7.2 \times 10^5}{4.0} \right)^{1/2} = 424.3 \text{ K} = 151^\circ\text{C} \quad (17.56)$$

Such isotope equilibration temperatures can be checked by determining the filling temperatures of fluid inclusions in sphalerite.

We now return to S-isotope fractionation in the oceans and to the variation of the $\delta^{34}\text{S}$ parameter in marine sulfate minerals throughout geological time. Marine sulfate in present-day seawater has $\delta^{34}\text{S} = +20.0\text{‰}$ throughout the oceans. When gypsum precipitates from sulfate-bearing solutions at Earth-surface temperatures, the solid is enriched in ^{34}S by only $\sim 1.65\text{‰}$. Therefore, marine gypsum and anhydrite effectively record the $\delta^{34}\text{S}$ value of sulfate ions in the oceans. Measurements of $\delta^{34}\text{S}$ values of marine gypsum/anhydrite deposits indicate that the $\delta^{34}\text{S}$ value of sulfate ions in the oceans has *varied systematically* throughout Phanerozoic and Precambrian time.

Table 17.5 Sulfur Isotope Fractionation Between Sulfide Minerals That Equilibrated Sulfur with the Same Reservoir at Elevated Temperatures in Kelvins

	Reference ^a
Pyrite–galena	
$10^3 \ln \alpha = (1.1 \times 10^6)/T^2$	(1)
Pyrite–sphalerite (or pyrrhotite)	
$10^3 \ln \alpha = (3.0 \times 10^5)/T^2$	(1)
Pyrite–chalcopyrite	
$10^3 \ln \alpha = (4.5 \times 10^5)/T^2$	(1)
Chalcopyrite–galena	
$10^3 \ln \alpha = (6.5 \times 10^5)/T^2$	(1)
Sphalerite–chalcopyrite and pyrrhotite–chalcopyrite	
$10^3 \ln \alpha = (1.5 \times 10^5)/T^2$	(1)
Sphalerite–galena	
$10^3 \ln \alpha = (8.9 \times 10^5)/T^2 - 0.57$	(2)
$10^3 \ln \alpha = (8.0 \times 10^5)/T^2$	(1)
$10^3 \ln \alpha = (7.0 \times 10^5)/T^2$	(3)
$10^3 \ln \alpha = (6.6 \times 10^5)/T^2 - 0.1$	(4)
Molybdenite–galena (experimental)	
$10^3 \ln \alpha = (1.3 \times 10^6)/T^2 - 0.80$	(5)
Molybdenite–sphalerite (experimental)	
$10^3 \ln \alpha = (0.71 \times 10^6)/T^2 - 0.15$	(5)
Molybdenite–pyrite (calculated, 400–750°C)	
$10^3 \ln \alpha = -(0.48 \times 10^6)/T^2 - 0.75$	(5)
Molybdenite–chalcopyrite (calculated, 400–750°C)	
$10^3 \ln \alpha = (0.72 \times 10^6)/T^2 - 0.70$	(5)
Molybdenite–pyrrhotite (calculated, 400–750°C)	
$10^3 \ln \alpha = (0.65 \times 10^6)/T^2 - 1.65$	(5)
SO ₂ (g)–H ₂ S(g) (287–1000°C)	
$10^3 \ln \alpha = (4.54 \times 10^6)/T^2 - 0.30$ (experimental)	(6)
$10^3 \ln \alpha = (3.65 \times 10^6)/T^2$ (calculated)	(6)
HSO ₄ [−] (aq)–H ₂ S(g) (200–320°C)	
$10^3 \ln \alpha = (5.1 \times 10^6)/T^2 + 6.3$	(7)
Sphalerite–HS [−] (aq) (50–340°C)	
$10^3 \ln \alpha = (1.11 \times 10^5)/T^2 + 1.36$	(2)
Galena–HS [−] (aq) (50–340°C)	
$10^3 \ln \alpha = (7.82 \times 10^5)/T^2 + 1.7$	(2)

^a(1) Kajiwaru and Krouse (1971); (2) Kiyosu (1973); (3) Czamanske and Rye (1974); (4) Grootenboer and Schwarcz (1969); (5) Suvorova (1974); (6) Thode et al. (1971); (7) Robinson (1973).

SOURCE: Friedman and O'Neil (1977).

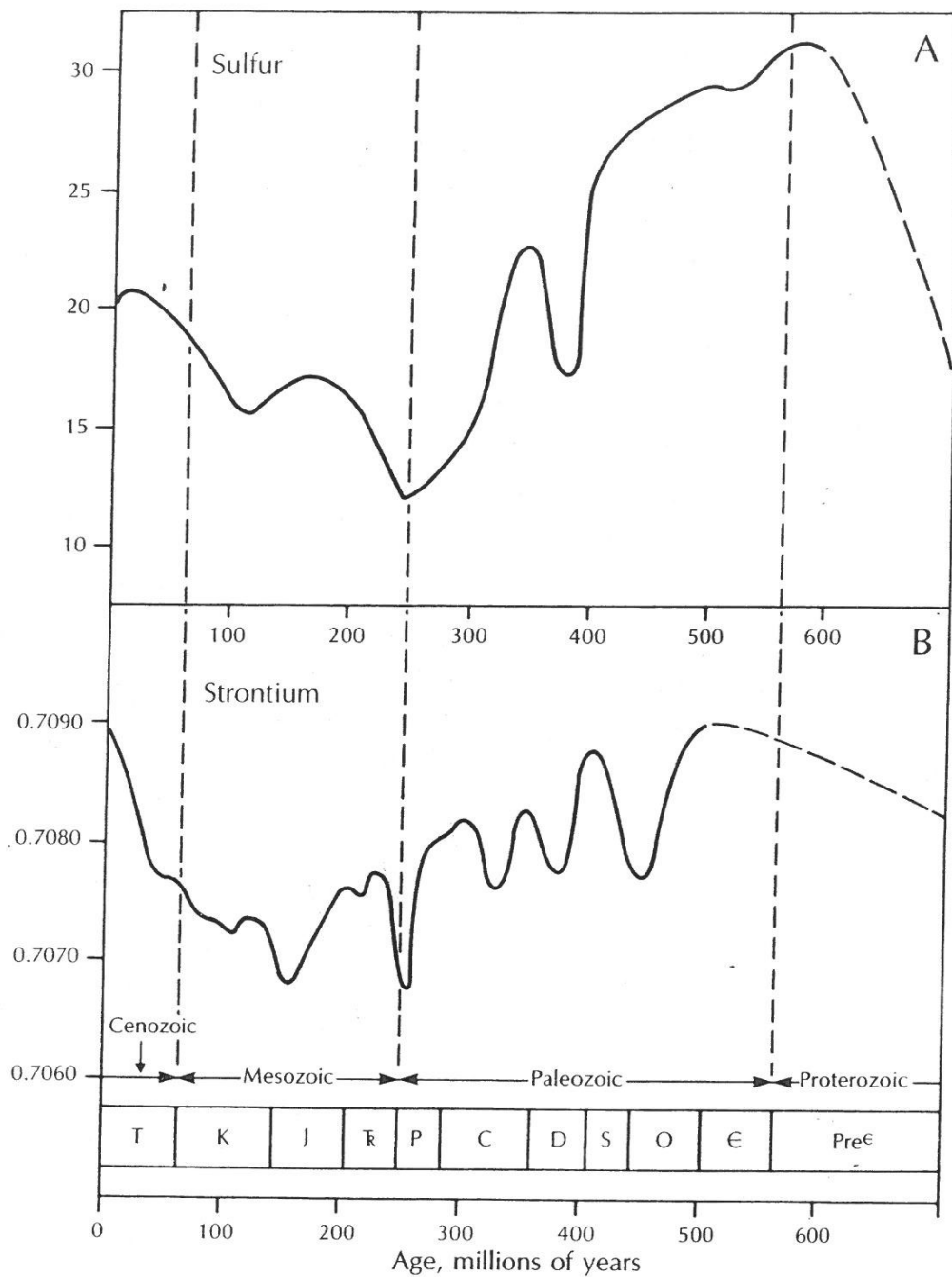


Figure 17.6 Variation of isotope compositions of S and Sr in the oceans in Phanerozoic time. Note that both $\delta^{34}\text{S}$ and the $^{87}\text{Sr}/^{86}\text{Sr}$ ratio declined irregularly during the Paleozoic Era from the Cambrian to the Permian periods. During the Mesozoic Era both parameters had low values but began to rise during the Cretaceous Period and continued to rise through the Tertiary Period. The observed variations of the isotopic compositions of S and Sr in the oceans were caused by global changes in the geochemical cycles of these elements. The curves for S and Sr are based on data by Claypool et al. (1980) and Burke et al. (1982), respectively. These references should be consulted for precise comparisons of the isotopic data pertaining to specific epochs or periods.

The isotope evolution of S in the oceans is remarkably similar to the isotopic evolution of Sr. A comparison of the two curves in Figure 17.6 shows that both the $\delta^{34}\text{S}$ value and the $^{87}\text{S}/^{86}\text{Sr}$ ratio in the oceans declined during the Paleozoic Era from high values during the Cambrian Period to low values during the Permian Period. Subsequently, both parameters remained low during the Mesozoic Era, but began to increase during the Cretaceous Period. The $^{87}\text{Sr}/^{86}\text{Sr}$ ratio has continued to increase steadily toward the Present (Figure 17.5), whereas the rise of the $\delta^{34}\text{S}$ parameter has faltered and may have been reversed in Late Tertiary time.

The isotope compositions of both elements have varied in response to changes in the magnitudes of Sr and S fluxes entering and leaving the oceans as well as in response to changes in the respective isotope compositions of the inputs and outputs. Sulfur enters the oceans primarily as the sulfate ion derived by erosion of sulfate deposits on the continents and by oxidation of sulfide minerals. In addition, S is discharged in hydrothermal fluids along spreading ridges on the ocean floor and by direct interactions between heated seawater and volcanic rocks of the oceanic crust. Sulfur is removed from the oceans by precipitation as insoluble metal sulfides and by direct precipitation of gypsum from seawater.

The isotope composition of S in the oceans is controlled by the $\delta^{34}\text{S}$ values of the inputs and outputs. For example, sulfate ions formed by oxidation of sedimentary sulfides in shale and limestone are depleted in ^{34}S and have negative $\delta^{34}\text{S}$ values, whereas sulfate ions released by erosion of marine evaporite rocks have positive $\delta^{34}\text{S}$ values. Therefore, the isotope composition of S entering the oceans from the continents at a particular time depends on the amounts of S each source contributes to the oceans. The output of S from the oceans by bacterial reduction of sulfate enriches the remaining sulfate in ^{34}S , whereas precipitation of gypsum lowers the S content of the oceans, making the isotope composition of the remaining S more susceptible to

change. Evidently, the $\delta^{34}\text{S}$ value of the oceans as well as its $^{87}\text{Sr}/^{86}\text{Sr}$ reflect the operation of the S and Sr cycles, respectively, on the surface of the Earth.

17.10 Summary

Isotope fractionation is a common phenomenon that affects the isotope compositions of H, C, O, N, S, and of a few other elements of low atomic number. It takes place during chemical reactions and in the course of certain physical processes, such as evaporation, condensation, and diffusion, because of differences in the masses of the stable isotopes of these elements. Isotope fractionation during chemical reactions results from the fact that the heavy isotope is concentrated preferentially in a particular compound during isotope exchange reactions that have achieved equilibrium at a specific temperature.

Isotope fractionation also takes place during unidirectional reactions because the masses of isotopes influence the rates of chemical reactions owing to differences in the bond strengths of isotopic molecules of the same compound or because of differences in the diffusion rates or vapor pressures of those isotopic molecules. Kinetic isotope fractionation effects are variable depending on environmental conditions and decrease as the reaction approaches equilibrium.

The isotope compositions of O and H in water molecules may vary widely because of isotope fractionation in the hydrologic cycle. Water vapor in the atmosphere is initially depleted in ^{18}O and D during the evaporation of seawater. Subsequently, the vapor is further depleted in ^{18}O and D because of the preferential removal of these isotopes by condensation of water or snow from moist air masses drifting from the equatorial region toward higher latitudes. Consequently, meteoric water is depleted in ^{18}O and D and has negative $\delta^{18}\text{O}$ and δD values on the SMOW scale, which are related by the equation for the so-called meteoric-water line.

The extent of isotope fractionation of meteoric water is generally related to latitude and temperature and hence to climatic conditions. The ice sheets of Greenland and Antarctica have preserved a record of climate change that can be recovered by systematic measurements of the isotope compositions of H and O in cores drilled through the ice sheets.

When calcium carbonate precipitates from aqueous solution, either inorganically or by the action of organisms living in the water, the O in the solid phase is *enriched* in ^{18}O compared to the water. When isotopic equilibrium is maintained, the isotope composition of O in the calcite depends on the temperature and on the isotope composition of O in the water. This relationship can be used to detect changes in the temperature of the oceans in the geologic past, provided that the isotope composition of the water has remained constant.

When clay minerals form by decomposition of aluminosilicate minerals, the isotope compositions of O and H in the clay are controlled by the *water* at the site of the reactions. Clay minerals in soil are enriched in ^{18}O but depleted in D compared to the meteoric water with which they are in contact. Their $\delta^{18}\text{O}$ and δD parameters are therefore related by linear equations having similar slopes but different intercept values than the meteoric-water line.

The O in groundwater is slowly *enriched* in ^{18}O by exchange with the O of silicate or carbonate minerals. The H-isotope composition of water may also change depending on the H-content of the rocks and on the presence of hydrocarbons or H_2S . The change in the isotope compositions increases with the temperature and salinity of the water and hence with depth below the surface. Seawater, trapped in the pore spaces of marine sediment at the time of deposition, is also altered by isotope exchange reactions with the rocks with which it comes in contact.

Carbon in the biosphere is depleted in ^{13}C during photosynthesis reactions depending on the

type of metabolism. Consequently, fossil fuels (coal, petroleum, and natural gas) are characteristically depleted in ^{13}C compared to the PDB standard. In addition, C-isotopes are fractionated by exchange reactions among CO_2 gas, carbonate ions in solution, and solid Ca carbonate. The process tends to enrich calcite or aragonite in ^{13}C relative to CO_2 gas of the atmosphere. However, when part or all of the CO_2 gas in an environment formed by oxidation of biogenic C compounds, the carbonate minerals inherit the biogenic C. Therefore, calcite cements in sandstones associated with petroleum pools and calcites in the caps of salt domes or associated with sedimentary U deposits may have negative $\delta^{13}\text{C}$ values on the PDB scale.

The isotope composition of Sr in marine carbonate rocks is *not* affected by isotope fractionation, but has nevertheless varied throughout geologic time as a result of mixing of different isotopic varieties of Sr entering the oceans. The isotope composition of Sr in stratigraphic sections of nonmarine carbonates may also vary in response to changes in the geology or hydrology of the drainage basin. The steady rise of the $^{87}\text{Sr}/^{86}\text{Sr}$ ratio of marine carbonates since the end of the Eocene epoch can be used to date carbonate rocks or fossil shells of Tertiary age.

The isotope composition of S is changed during the reduction of sulfate to sulfide by anaerobic bacteria. As a result, the remaining sulfate is enriched in ^{34}S , whereas the sulfide is enriched in ^{32}S and depleted in ^{34}S . Therefore, sulfide minerals in unmetamorphosed sedimentary rocks commonly have negative $\delta^{34}\text{S}$ values relative to troilite of Canyon Diablo whereas marine sulfate rocks have positive $\delta^{34}\text{S}$ values. The isotope composition of S in marine sulfate rocks of Precambrian and Phanerozoic age has varied systematically with time. The time-dependent S-isotope curve in the oceans is similar to the Sr-isotope curve, which suggests that both elements were affected by global geological changes that modified the cycles of Sr and S in similar ways.

Problems

1. Calculate $\delta^{18}\text{O}$ of water *vapor* in equilibrium with liquid water at 10°C assuming that $\delta^{18}\text{O}_\text{l} = -10.0\text{‰}$ (SMOW) and $\alpha_\text{v}^1 = 1.0105$. (Answer: -20.3‰)
2. Calculate $\delta^{18}\text{O}$ of liquid water that is condensing from vapor at 10°C having $\delta^{18}\text{O} = -25.0\text{‰}$ assuming that $\alpha_\text{v}^1 = 1.0105$. (Answer: -14.8‰)
3. Calculate the value of α_v^1 at 35°C from equation 17.34. (Answer: 1.008 43)
4. Calculate the value of δD of water *vapor* remaining in an air mass based on Rayleigh distillation for the case that $\alpha_\text{v}^1(\text{D}) = 1.074$, $f = 0.45$, $\delta\text{D}_\text{v}^0 = -94.8\text{‰}$. (Answer: -146.7‰)
5. Calculate the value of δD of *liquid* water condensing from an air mass whose initial $\delta\text{D} = -94.8\text{‰}$, if only 10% of the original vapor remain and $\alpha_\text{v}^1(\text{D}) = 1.074$. (Answer: -180.1‰)
6. Calculate the $\delta^{18}\text{O}$ value of meteoric water whose $\delta\text{D} = -70\text{‰}$. (Answer: -10‰)
7. What is the $\delta^{18}\text{O}$ value of calcite precipitated at 5°C in equilibrium with meteoric water having $\delta\text{D} = -85\text{‰}$? (Answer: $+22.7\text{‰}$)
8. Given that a fluid inclusion filled with water in calcite that formed in a cave has a measured value of $\delta\text{D} = -175\text{‰}$ and that $\delta^{18}\text{O}$ of the calcite is $+10.0\text{‰}$ (SMOW). What was the temperature of deposition of the calcite in $^\circ\text{C}$ assuming that the water was of meteoric origin and isotopic equilibrium existed between calcite and water at the time of depositions? (Answer: 3.8°C)
9. What is the $\delta^{18}\text{O}$ value of water on the SMOW scale in a fluid inclusion in calcite whose $\delta^{18}\text{O} = -2.5\text{‰}$ (PDB), if the O in the water reequilibrated with the O of the calcite at 15°C after deposition? (Answer: -2.70‰)
10. Compare Problems 8 and 9 and explain why the measured $\delta^{18}\text{O}$ values of water in fluid inclusions in calcite or quartz cannot be used to determine the temperature of formation of these minerals.
11. Calculate $\delta^{18}\text{O}$ and δD of glauconite in equilibrium with seawater ($\delta^{18}\text{O} = 0\text{‰}$, $\delta\text{D} = 0\text{‰}$) using the fractionation factors in Table 17.2. (Answer: $\delta^{18}\text{O} = +26\text{‰}$, $\delta\text{D} = -70\text{‰}$)
12. Calculate $\alpha_\text{w}^q(\text{O})$ for the quartz–water at temperatures of 20, 40, 60, 80, and 100°C and plot the results in coordinates of α_w^q and $t^\circ\text{C}$. (Answer: 1.036 58 at 20°C , 1.031 55 at 40°C , 1.027 42 at 60°C , 1.023 98 at 80°C , and 1.021 09 at 100°C)
13. What is the $\delta^{13}\text{C}$ (PDB) value of calcite precipitated at 15°C in equilibrium with CO_2 gas having $\delta^{13}\text{C} = -25.0\text{‰}$ (PDB)? (Answer: -14.5‰ (PDB))
14. Sulfate in a reservoir of limited size is being reduced by bacteria at 10°C and the resulting H_2S gas escapes from the system. Assume that Rayleigh distillation is occurring and plot $\delta^{34}\text{S}$ trajectories both for the remaining sulfate and for the H_2S gas for values of f from 1.0 to 0.01. Assume that $\alpha_{\text{H}_2\text{S}}^{\text{SO}_4} = 1.025$ and that $\delta^{34}\text{S}^0(\text{sulfate}) = +20\text{‰}$.

References

- ANDERSON, T. F., and M. A. ARTHUR, 1983. Stable isotopes of oxygen and carbon and their application to sedimentologic and paleoenvironmental problems. In M. A. Arthur, T. F. Anderson, I. R. Kaplan, J. Veizer, and L. S. Land (Eds.), *Stable Isotopes in Sedimentary Geology*. SEPM Short Course No. 10, Soc. Econ. Paleont. Mineral., Tulsa, OK.
- ARTHUR, M. A., T. F. ANDERSON, I. R. KAPLAN, J. VEIZER, and L. S. LAND (Eds.), 1983. *Stable Isotopes in Sedimentary Geology*. SEPM Short Course No. 10, Soc. Econ. Paleont. Mineral., Tulsa, OK.
- BIGEISEN, J., 1965. Chemistry of isotopes. *Science*, 147:463–471.

- BURKE, W. H., R. E. DENISON, E. A. HETHERINGTON, R. B. KOEPNICK, N. F. NELSON, and J. B. OTTO, 1982. Variation of seawater $^{87}\text{Sr}/^{86}\text{Sr}$ throughout Phanerozoic time. *Geology*, **10**:516–519.
- CHENEY, E. S., and M. L. JENSEN, 1967. Corrections to carbon isotope data of Gulf Coast salt-dome cap rock. *Geochim. Cosmochim. Acta*, **31**:1345–1346.
- CLAYTON, R. N., I. FRIEDMAN, D. L. GRAF, T. K. MAYEDA, W. F. MEENTS, and N. F. SHIMP, 1966. The origin of saline formation waters. I. Isotopic composition. *J. Geophys. Res.*, **71**(16):3869–3882.
- CLAYPOOL, G. E., W. T. HOLSER, I. R. KAPLAN, H. SAKAI, and I. ZAK, 1980. The age curves of sulfur and oxygen isotopes in marine sulfate and their mutual interpretation. *Chem. Geol.*, **28**:199–260.
- CZAMANSKE, G. K., and R. O. RYE, 1974. Experimentally determined sulfur isotope fractionations between sphalerite and galena in the temperature range 600° to 275°C. *Econ. Geol.*, **69**:17–25.
- DEINES, P., D. LANGMUIR, and R. S. HARMON, 1974. Stable carbon isotope ratios and the existence of a gas phase in the evolution of carbonate groundwater. *Geochim. Cosmochim. Acta*, **38**:1147–1164.
- DEINES, P., 1980. The isotopic composition of reduced organic carbon. In P. Fritz and J. Ch. Fontes (Eds.), *Handbook of Environmental Isotope Geochemistry*, vol. 1, 329–406. Elsevier, Amsterdam, 545 pp.
- DEPAOLO, D. J., and B. L. INGRAM, 1985. High-resolution stratigraphy with strontium isotopes. *Science*, **227**:938–941.
- DING, T., S. JIANG, D. WAN, Y. LI, J. LI, H. SONG, Z. LIU, and X. YAO, 1996. *Silicon Isotope Geochemistry*. Geol. Publ. House, Beijing, P. R. C., 125 pp.
- DONOVAN, T. J., I. FRIEDMAN, and J. D. GLEASON, 1974. Recognition of petroleum-bearing traps by unusual isotopic compositions of carbonate-cemented surface rocks. *Geology*, **2**:351–354.
- FAURE, G. 1982. The marine-strontium geochronometer. In G. S. Odin (Ed.), *Numerical Dating in Stratigraphy*, vol. 1, 73–79. Wiley, Chichester, England, 630 pp.
- FAURE, G., 1986. *Principles of Isotope Geology*, 2nd ed. Wiley, New York, 589 pp.
- FRIEDMAN, I., and J. R. O'NEIL, 1977. Compilation of stable isotope fractionation factors of geochemical interest. In M. Fleischer (Ed.), *Data of Geochemistry*, 6th ed., ch. kk. U. S. Geol. Surv. Prof. Paper 440.
- FRITZ, P., and J. CH. FONTES (Eds.), 1980a. *Handbook of Environmental Isotope Geochemistry*, vol. 1. Elsevier, Amsterdam, 545 pp.
- FRITZ, P., and J. CH. FONTES, 1980b. Introduction. In P. Fritz and J. Ch. Fontes (Eds.), *Handbook of Environmental Isotope Geochemistry*, vol. 1, 1–18. Elsevier, Amsterdam, 545 pp.
- FRITZ, P., and J. CH. FONTES (Eds.), 1986. *Handbook of Environmental Isotope Geochemistry*, vol. 2. Elsevier, Amsterdam, 557 pp.
- FRY, B., J. COX, H. GEST, and J. M. HAYES, 1986. Discrimination between ^{34}S and ^{32}S during bacterial metabolism of inorganic sulfur compounds. *J. Bacteriol.*, **165**:328–330.
- GROOTENBOER, J., and H. P. SCHWARCZ, 1969. Experimentally determined sulfur isotope fractionations between sulfide minerals. *Earth Planet. Sci. Lett.*, **7**:162–166.
- HITCHON, B., and I. FRIEDMAN, 1969. Geochemistry and origin of formation waters in the western Canada sedimentary basin—I. Stable isotopes of hydrogen and oxygen. *Geochim. Cosmochim. Acta*, **33**:1321–1349.
- HOEFS, J., 1996. *Stable Isotope Geochemistry*, 4th ed. Springer-Verlag, Berlin, 241 pp.
- HURST, R. W., 1986. Strontium isotope chronostratigraphy and correlation of the Miocene Monterey Formation in the Ventura and Santa Maria basins of California. *Geology*, **14**:459–462.
- KAJIWARA, Y., and H. R. KROUSE, 1971. Sulfur isotope partitioning in metallic sulfide systems. *Can. J. Earth Sci.*, **8**:1397–1408.
- KEITH, M. L., and J. N. WEBER, 1964. Carbon and oxygen isotopic composition of selected limestones and fossils. *Geochim. Cosmochim. Acta*, **28**:1787–1816.
- KHARAKA, Y. K., and W. W. CAROTHERS, 1986. Oxygen and hydrogen isotope geochemistry of deep basin brines. In P. Fritz and J. Ch. Fontes (Eds.), *Handbook of Environmental Isotope Geochemistry*, vol. 2, 305–360. Elsevier, Amsterdam, 557 pp.
- KIYOSU, Y., 1973. Sulfur isotope fractionation among sphalerite, galena and sulfide ions. *Geochem. J.*, **7**:191–199.
- JOHNSON, S., J. W. DANSGAARD, H. B. CLAUSEN, and C. C. LANGWAY, JR., 1972. Oxygen isotope profiles through the Antarctic and Greenland ice sheets. *Nature*, **235**:429–434.
- LAWRENCE, J. R., and H. P. TAYLOR, 1971. Deuterium and oxygen-18 correlation: clay minerals and hydroxides in Quaternary soils compared to meteoric waters. *Geochim. Cosmochim. Acta*, **35**:993–1003.
- LORD, B. K., L. M. JONES, and G. FAURE, 1988. Evidence for the existence of the Gondwana ice sheet in the ^{18}O depletion of carbonate rocks in the Permian formations of the Transantarctic Mountains. *Chem. Geol., Isotope Geosci. Sect.*, **8**:163–172.
- LORIUS, C., J. JOUZEL, C. RITZ, L. MERLIVAT, N. I. BARKOV, Y. S. KOROTKEVICH, and V. M. KOTLYAKOV, 1985. A 150,000-year climatic record from Antarctic ice. *Nature*, **316**:591–596.
- LOWRY, R. M., G. FAURE, D. I. MULLET, and L. M. JONES, 1988. Interpretation of chemical and isotopic compositions of brines based on mixing and dilution. "Clinton" sandstones, eastern Ohio, U.S.A. *Appl. Geochem.*, **3**:177–184.

- MOSER, H., and W. STICHLER, 1981. Environmental isotopes in ice and snow. In P. Fritz and J. Ch. Fontes (Eds.), *Handbook of Environmental Isotope Geochemistry*, vol. 1, 141-178. Elsevier, Amsterdam, 545 pp.
- NEAT, P. L., G. FAURE, and W. J. PEGRAM, 1979. The isotopic composition of strontium in non-marine carbonate rocks: The Flagstaff Formation of Utah. *Sedimentology*, **26**:271-282.
- OHMOTO, H., and R. O. RYE, 1979. Isotopes of sulfur and carbon. In H. L. Barnes (Ed.), *Geochemistry of Hydrothermal Ore Deposits*, 2nd ed., 509-567. Wiley, New York, 798 pp.
- O'NEIL, J. R., and Y. K. KHARAKA, 1976. Hydrogen and oxygen isotope exchange reactions between clay minerals and water. *Geochim. Cosmochim. Acta*, **40**:241-246.
- ROBINSON, R., 1973. Sulfur isotope equilibrium during sulfur hydrolysis at high temperature. *Earth Planet. Sci. Lett.*, **18**:443-450.
- RUNDEL, P. W., J. R. EHLERINGER, and K. A. NAGY (Eds.), 1988. *Stable Isotopes in Ecological Research*. Springer-Verlag, New York, 570 pp.
- SAVIN, S. M., and S. EPSTEIN, 1970. The oxygen and hydrogen geochemistry of clay minerals. *Geochim. Cosmochim. Acta*, **34**:25-42.
- STANLEY, K. O., and G. FAURE, 1979. Isotopic composition and source of strontium in sandstone cements: The High Plains Sequence of Wyoming and Nebraska. *J. Sed. Pet.*, **49**:45-54.
- SUVOROVA, V. A., 1974. Temperature dependence of the distribution coefficient of sulfur isotopes between equilibrium sulfides. Fifth National Symp. on Stable Isotope Geochemistry, Moscow, Pt. 1, 128 pp.
- THODE, H. G., C. B. CRAGG, J. B. HULSTON, and C. E. REES, 1971. Sulfur isotope exchange between sulfur dioxide and hydrogen sulfide. *Geochim. Cosmochim. Acta*, **35**:35-45.
- VALLEY, J. W., H. P. TAYLOR, JR., and J. R. O'NEIL (eds.), 1986. Stable isotopes in high temperature geological processes. *Rev. Mineral.*, vol. 16. Mineral. Soc. Amer., Washington, DC, 570 pp.
- VEIZER, J., and J. HOEFS, 1967. The nature of $^{18}\text{O}/^{16}\text{O}$ and $^{13}\text{C}/^{12}\text{C}$ secular trends in sedimentary carbonate rocks. *Geochim. Cosmochim. Acta*, **40**:1387-1395.
- WEAST, R. C., M. J. ASTLE, and W. H. BEYER (Eds.), 1986. *CRC Handbook of Chemistry and Physics*. CRC Press, Boca Raton, FL.
- YEH, H. W., 1980. D/H ratios and late stage hydration of shales during burial. *Geochim. Cosmochim. Acta*, **45**:341-352.
- YURTSEVER, Y., 1975. Worldwide survey of stable isotopes in precipitation. Rept. Section Isotope Hydrology, International Atomic Energy Agency, Vienna, 40 pp.