

H 2.1																	He
Li 1.0	Be 1.5											B 2.0	C 2.5	N 3.0	O 3.5	F 4.0	Ne
Na 0.9	Mg 1.2											Al 1.5	Si 1.8	P 2.1	S 2.5	Cl 3.0	Ar
K 0.8	Ca 1.0	Sc 1.3	Ti 1.5	V 1.6	Cr 1.6	Mn 1.5	Fe 1.8	Co 1.9	Ni 1.9	Cu 1.9	Zn 1.6	Ga 1.6	Ge 1.8	As 2.0	Se 2.5	Br 2.8	K
Rb 0.8	Sr 1.0	Y 1.2	Zr 1.4	Nb 1.6	Mo 1.8		Ru 2.2	Rh 2.2	Pd 2.2	Ag 1.9	Cd 1.7	In 1.7	Sn 1.8	Sb 1.9	Te 2.1	I 2.5	Xe
Cs 0.7	Ba 0.9	La 1.0	Hf 1.3	Ta 1.5	W 1.7	Re 1.9	Os 2.2	Ir 2.2	Pt 2.2	Au 2.4	Hg 1.9	Tl 1.8	Pb 1.9	Bi 1.9	Po 2.2	At 2.2	Rd
Fr 0.7	Ra 0.9	Ac 1.1															
			La 1.0	Ce 1.0	Pr 1.0	Nd 1.0		Sm 1.0	Eu 1.1	Gd 1.1	Tb 1.1	Dy 1.1	Ho 1.1	Er 1.1	Tm 1.2	Yb 1.2	Lu 1.2
			Ac 1.1	Th 1.3	Pa 1.4	U 1.4											

Figure 1.5 Electronegativities of the elements.

reached. Thus Ne has little tendency to either add or give up an electron.

The number of electrons that an element will either give up or accept is known as its *valence*. For elements in the “wings” of the periodic table (i.e., all except the transition metals), valence is easily determined simply by counting how far the element is horizontally displaced from Group 18 in the periodic table. For Group 18, this is 0, so these elements, the noble gases, have 0 valence. For group 1 it is 1, so these elements have valence of +1; for group 17 it is -1, so these elements have valence of -1. Valence of the transition metals is not so simply determined, and these elements can have more than one valence state. Most, however, have valence of 2 or 3, though some, such as U, can have valences as high as 6.

A final characteristic that is important in controlling chemical properties is *ionic radius*. This is deduced from bond length when the atom is bonded to one or more other atoms. Positively charged atoms, or *cations*, have smaller ionic radii than do negatively charged atoms, or *anions*. Also, ionic radius decreases as charge increases. This decrease is due both to loss of outer electrons and to shrinking of the orbits of the remaining electrons. The latter occurs because the charge of the nucleus is shared by fewer electrons and hence has a greater attractive force on each. In addition,

ionic radius increases downward in each group in the periodic table, both because of addition of electrons to outer shells and because these outer electrons are increasingly shielded from the nuclear charge by the inner ones. Ionic radius is important in determining geochemical properties such as substitution in solids, solubility, and diffusion rates. Large ions need to be surrounded, or *coordinated*, by a greater number of oppositely charged ions than do smaller ones. The ionic radii of the elements are illustrated in Figure 1.6.

We can now summarize a few of the more important chemical properties of the various groups in the periodic table. Group 18 does not participate in chemical bonding in nature, hence the term “noble gases”. Group 1 elements, the alkalis, readily accept an electron (they are *electropositive*) and hence are highly reactive. They tend to form *ionic bonds* rather than *covalent* ones and hence tend not to form bonds that are as strong as other elements. They tend to be quite soluble in aqueous solutions. Because they have only a +1 charge, their ionic radii tend to be larger than those of other cations. Group 2 elements, the *alkaline earths*, have these same characteristics, but somewhat moderated. Group 17 elements, the *halogens*, are highly electronegative and readily accept an electron, are highly reactive, form ionic bonds, and are quite soluble. Their ionic radii tend to be

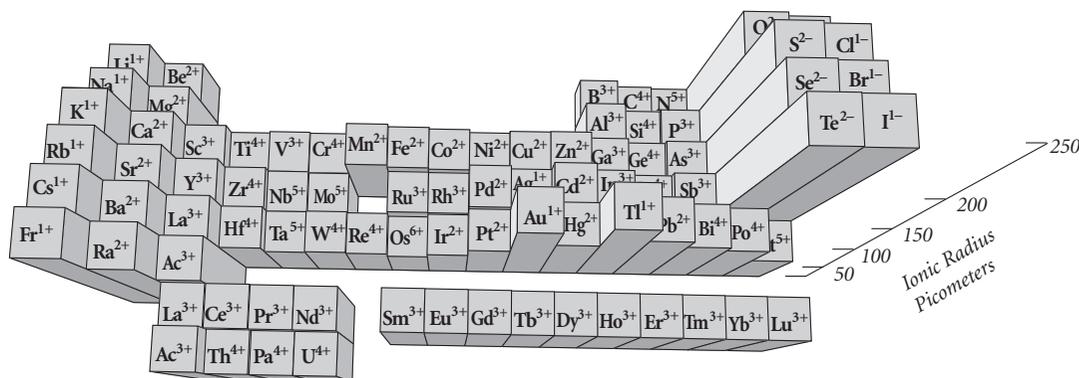


Figure 1.6 Ionic radii of the elements.

larger than more highly charged anions. Elements of Groups 13–16 tend to form bonds that are predominantly covalent. As a result, they tend to be less reactive and less soluble (except where they form soluble radicals, such as SO_4^{2-}) than Group 1, 2, and 17 elements. Finally, the transition metals are a varied lot. Many form strong bonds (generally with O in nature) and can have quite variable solubility (which, as we shall see in Chapter 6, depends on their valence state). Some, the “noble metals” (Ru, Rh, Pd, Os, Ir, Pt) in particular, tend to be very unreactive and insoluble. The rare earths are of interest because all have two electrons in the 6s outer orbital, and vary only by the number of electrons in the 4f shell. Because they almost always have the same valence (+3), bonding behavior is quite similar. They vary systematically in ionic radius, which results in a systematic variance in geochemical behavior, as we will see in Chapter 7.

1.4.4 Chemical bonding

1.4.4.1 Covalent, ionic, and metal bonds

Except for the noble gases, atoms rarely exist independently; they are generally bound to other atoms in molecules, crystals, or ionic radicals. Atoms bind to one another through transfer or sharing of electrons, or through electrostatic forces arising from uneven distribution of charge in atoms and molecules. A bond that results from the transfer of electrons from one atom to another is known as an *ionic bond*. A good example is the bond

between Na and Cl in a halite crystal. In this case, the Na atom (the electropositive element) gives up an electron, becoming positively charged, to the Cl atom (the electronegative element), which becomes negatively charged. Electrostatic forces between the Na^+ and the Cl^- ions hold the ions in place in the crystal. When electrons are shared between atoms, such as in the H_2O molecule or the SiO_4^{4-} radical, the bond is known as covalent. In a *covalent bond*, the outer electrons of the atoms involved are in hybrid orbits that encompass both atoms.

The ideal covalent and ionic bonds represent the extremes of a spectrum: most bonds are neither wholly covalent nor wholly ionic. In these intermediate cases, the bonding electrons will spend most, but not all, of their time associated with one atom or another. Electronegativity is useful in describing the degree of ionicity of a bond: a bond is considered ionic when the difference in the electronegativity of the two atoms involved is greater than 2. In Figure 1.5, we see that metals (generally those elements on the left-hand side of the periodic table) tend to have low electronegativities while the non-metals (those elements on the right) have high electronegativities. Thus bonds between metals and non-metals (e.g., NaCl) will be ionic while those between non-metals (e.g., CO_2) will be covalent, as will bonds between two like atoms (e.g., O_2).

Another type of bond occurs in pure metal and metal alloy solids. In the metallic bond, valence electrons are not associated with any single atom or pair of atoms; rather, they are

mobile and may be found at any site in the crystal lattice. Since metals rarely occur naturally at the surface of the Earth (they do occur in meteorites and the Earth's core), this type of bond is less important in geochemistry than other bonds.

Ionically bonded compounds tend to be good conductors of heat and electricity and have high melting temperatures. They also tend to be hard, brittle and highly soluble in water. Covalently bonded compounds tend to be good conductors of heat, but not of electricity. They are typically harder and less brittle than ionic solids but less soluble. In molecular solids, such as ice, atoms within the molecule are covalently bonded. The molecules themselves, which occupy the lattice points of the crystal, are bonded to each other through van der Waals and/or hydrogen bonds. Such solids are comparatively weak and soft and generally have low melting points.

Molecules in which electrons are unequally shared have an asymmetric distribution of charge and are termed *polar*. A good example is the hydrogen chloride molecule. The difference in electronegativity between hydrogen and chlorine is 0.9, so we can predict that the bonding electron will be shared, but associated more with the Cl atom than with the H atom in HCl. Thus the H atom will have a partial positive charge, and the Cl atom a partial negative charge. Such a molecule is said to be a *dipole*. The *dipole moment*, which is the product of one of the charges (the two charges are equal and opposite) times the distance between the charges, is a measure of the asymmetric distribution of charge. Dipole moment is usually expressed in debye units (1 debye = 3.3356×10^{-34} coulomb-meters).

1.4.4.2 Van der Waals interactions and hydrogen bonds

Covalent and ionic bonds account for the majority of bonds between atoms in molecules and crystals. There are two other interactions that play a lesser role in interactions between atoms and molecules: van der Waals interactions and hydrogen bonds. These are much weaker but nevertheless play an important role in chemical interactions, particularly where water and organic substances are involved.

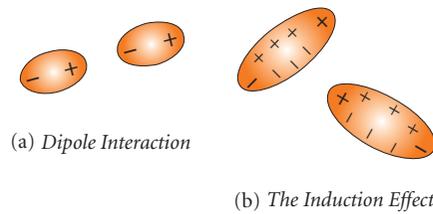


Figure 1.7 Van der Waals interactions arise because of the polar nature of some molecules. Illustrated here are (a) dipole–dipole interactions, which occur when two dipolar molecules orient themselves so that oppositely charged sides are closest, and (b) the induction effect, which arises when the electron orbits of one molecule are perturbed by the electromagnetic field of another molecule.

Van der Waals interactions arise from asymmetric distribution of charge in molecules and crystals. There are three sources for van der Waals interactions: *dipole–dipole attraction*, *induction*, and *London dispersion* forces. As we noted earlier, many molecules, including water, have permanent dipole moments. When two polar molecules encounter each other, they will behave much as two bar magnets: they will tend to orient themselves so that the positive part of one molecule is closest to the negative part of another (Figure 1.7a). This results in a net attractive force between the two molecules. When the distance between molecules is large compared with the distance between charges within molecules, the energy of attraction can be shown to be:

$$U_{D-D} = -\frac{2}{3} \frac{\mu^4}{r^6} \frac{1}{kT} \quad (1.3)$$

where U_{D-D} is the interaction energy, μ is the dipole moment, T is temperature (absolute, or thermodynamic temperature, which we will introduce in the next chapter), k is a constant (Boltzmann's constant, which we shall also meet in the next chapter), and r is distance. We do not want to get lost in equations at this point; however, we can infer several important things about dipole–dipole interactions just from a quick glance at it. First, the interaction energy depends inversely on the sixth

power of distance. Many important forces, such as electromagnetic and gravitational forces, depend on the inverse square of distance. Thus we may infer that dipole–dipole forces become weaker with distance very rapidly. Indeed, they are likely to be negligible unless the molecules are very close. Second, the interaction energy depends on the fourth power of the dipole moment, so that small differences in dipole moment will result in large differences in interaction energy. For example, the dipole moment of water (1.84 debyes) is less than twice that of HCl (1.03 debyes), yet the dipole interaction energy between two water molecules (716 J/mol) is nearly 10 times as great as that between two HCl molecules (72.24 J/mol) at the same temperature and distance (298 K and 50 pm). Finally, we see that dipole interaction energy will decrease with temperature.

Dipole molecules may also polarize electrons in a neighboring molecule and distort their orbits in such a way that their interaction with the dipole of the first molecule is attractive. This is known as the induction effect (Figure 1.7b). The induction energy also depends on the inverse sixth power of intermolecular distance, but only on the square of the dipole moment of the molecules involved. In addition, another parameter, the polarizability of a molecule, is also needed to describe this effect. In general, the attraction arising from induction is less important than from dipole–dipole interaction. However, because it depends only on the square of dipole moment, the induction attraction can be larger than the dipole–dipole attraction for some weakly dipolar molecules.

Finally, van der Waals forces can also occur as a consequence of fluctuations of charge distribution on molecules that occur on time-scales of 10^{-16} seconds. These are known as London dispersion forces. They arise when the instantaneous dipole of one molecule induces a dipole in a neighboring molecule. As was the case in induction, the molecules will orient themselves so that the net forces between them are attractive.

The total energy of all three types of van der Waals interactions between water molecules is about 380 J/mol assuming an intermolecular distance of 5 pm and a temperature of 298 K (25°C). Though some interaction energies can be much stronger (e.g., CCl_4 , 2.8 kJ/mol)

or weaker (1 J/mol for He), an energy of a few hundred joules per mole is typical of many substances. By comparison, the hydrogen–oxygen bond energy for each H–O bond in the water molecule is 46.5 kJ/mol. Thus van der Waals interactions are quite weak compared with typical intramolecular bond energies.

The *hydrogen bond* is similar to van der Waals interactions in that it arises from non-symmetric distribution of charge in molecules. However, it differs from van der Waals interactions in a number of ways. First, it occurs exclusively between hydrogen and strongly electronegative atoms, namely oxygen, nitrogen and fluorine. Second, it can be several orders of magnitude stronger than van der Waals interactions, though still weak by comparison with covalent and ionic bonds. The exact nature of hydrogen bonds is not completely understood. They arise principally from electrostatic interactions. In the water molecule, bonding between oxygen and hydrogen results in hybridization of s and p orbitals to yield two bonding orbitals between the O and two H atoms, and two non-binding sp^3 orbitals on the oxygen. The latter are prominent on the opposite side of the O from the hydrogens. The hydrogen in one water molecule, carrying a net positive charge, is attracted by the non-binding sp^3 electrons of the oxygen of another water molecule, forming a hydrogen bond with it (Figure 1.8).

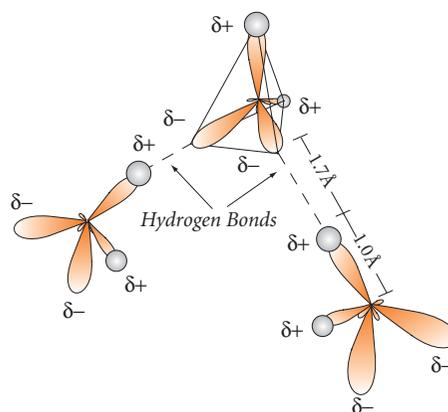


Figure 1.8 Hydrogen bonding between water molecules. Hydrogen positions are shown as red; sp^3 hybrid orbitals in oxygen are shown as dark gray. The δ^+ and δ^- indicate partial positive and negative charges respectively.

Hydrogen bonds typically have energies in the range of 20–40 kJ/mol. These are much higher than expected for electrostatic interactions alone, and indeed approach values similar to intramolecular bond energies. Thus there is the suspicion that to some degree of covalency is also involved in the hydrogen bond. That is to say, the non-binding electrons of oxygen are to some degree shared with the hydrogen in another molecule. Hydrogen bonds are perhaps most important in water, where they account for some of the extremely usual properties of this compound, such as its high heat of vaporization, but they can also be important in organic molecules and are present in HF and ammonia as well.

1.5 A BRIEF LOOK AT THE EARTH

1.5.1 Structure of the Earth

The Earth consists of three principal layers: the core, the mantle, and the crust (Figure 1.9). The core, roughly 3400 km thick and extending about halfway to the surface, consists of Fe–Ni alloy and can be subdivided into an inner and outer core. The outer core is liquid while the inner core is solid. The mantle is about 3000 km thick and accounts for about two-thirds of the mass of the Earth,

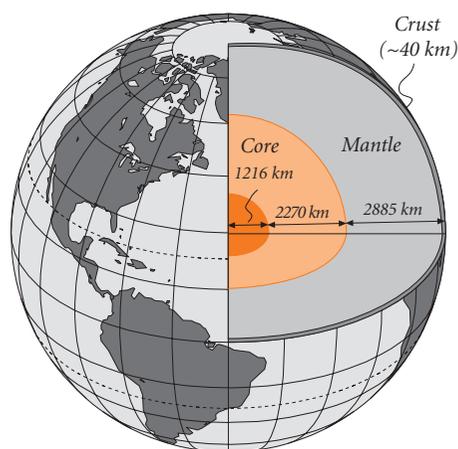


Figure 1.9 The Earth in cross-section. The outer rocky part of the planet, the mantle and crust, consists principally of silicates and is 2885 km thick. The core, divided into a liquid outer core and a solid inner core, consists of iron–nickel alloy and is 3486 km thick.

with the core accounting for the other third. The crust is quite thin by comparison, nowhere thicker than 100 km and usually much thinner. Its mass is only about 0.5% of the mass of the Earth. There are two fundamental kinds of crust: oceanic and continental. Ocean crust is thin (about 6 km) and, with the exception of the eastern Mediterranean, is nowhere older than about 180 million years. The continental crust is thicker (about 35–40 km thick on average) and relatively permanent, with an average age of 2 billion years.

Both the crust and the mantle consist principally of silicates. The mantle is comparatively rich in iron and magnesium, so ferromagnesian silicates, such as olivine and pyroxenes, dominate. Rocks having these compositional characteristics are sometimes called *ultramafic*. The continental crust is poor in iron and magnesium, and aluminosilicates such as feldspars dominate. Rocks of this composition are sometimes referred to as *felsic*. The oceanic crust is intermediate in composition between the mantle and continental crust and has a *mafic* composition, consisting of a roughly 50:50 mix of ferromagnesian minerals and feldspar. These differences in composition lead to differences in density, which are ultimately responsible for the layering of the Earth, the density of each layer decreasing outward. The continental crust is the least dense of these layers. The fundamental reason why continents stick out above the oceans is that continental crust is less dense than oceanic crust.

1.5.2 Plate tectonics and the hydrologic cycle

Two sources of energy drive all geologic processes: solar energy and the Earth's internal heat. Solar energy drives atmospheric and oceanic circulation, and with them, the hydrologic cycle. In the hydrologic cycle, water vapor in the atmosphere precipitates on the land as rain or snow, percolates into the soil and, through the action of gravity, makes its way to the oceans. From the oceans, it is evaporated into the atmosphere again and the cycle continues. The hydrologic cycle is responsible for two very important geologic processes: weathering and erosion. Weathering causes rocks to break down into small particles and dissolved components. The particles and dissolved matter are carried by the

flow of water (and more rarely by wind and ice) from high elevation to areas of low elevation. Thus the effect of the hydrologic cycle is to level the surface of the planet.

The Earth's internal heat is responsible for tectonic processes, which deform the surface of the planet, producing topographic highs and lows. The internal heat has two parts. Some fraction of the heat, 50% or more, originated from the gravitational energy released when the Earth formed. The other fraction of internal heat is produced by the decay of radioactive elements, principally uranium, thorium, and potassium, in the Earth. The Earth's internal heat slowly decays over geologic time as it migrates to the surface and is radiated away into space. It is this migration of heat out of the Earth that drives tectonic processes. Heat causes both the outer core and the mantle to convect, as hot regions rise and cold regions sink. Convection within the outer core gives rise to the Earth's magnetic field, and may have other, as yet not understood, geologic consequences. Convection in the mantle is responsible for deformation of the Earth's crust as well as volcanism.

The great revolution in earth science in the 1960s centered on the realization that the outer part of the Earth was divided into a number of "plates" that moved relative to one another. Most tectonic processes, as well as most volcanism, occur at the boundaries between these plates. The outer part of the Earth, roughly the outer 100 km or so, is cool enough ($<1000^{\circ}\text{C}$) that it is rigid. This rigid outer layer is known as the lithosphere and comprises both the crust and the outermost mantle (Figure 1.10). The mantle below the lithosphere is hot enough (and under sufficient confining pressure) that it flows, albeit extremely slowly, when stressed. This part of the mantle is known as the asthenosphere. Temperature differences in the mantle create stresses that produce convective flow. It is this flow that drives the motion of the lithospheric plates. The motion of the plates is extremely slow, a few tens of centimeters per year at most and generally much less. Nevertheless, on geologic time-scales they are sufficient to continually reshape the surface of the Earth, creating the Atlantic Ocean, for example, in the last 200 million years.

Rather than thinking of plate motion as being driven by mantle convection, it would

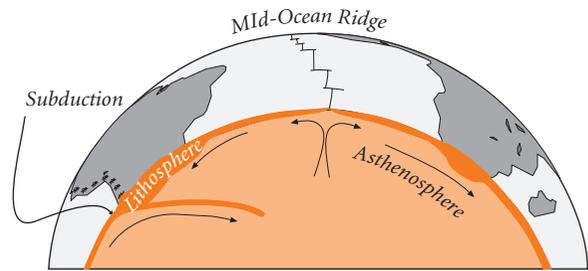


Figure 1.10 Cross-section of the Earth illustrating relationships between lithosphere and asthenosphere and plate tectonic processes. Oceanic crust and lithosphere are created as plates diverge at mid-ocean ridges and are eventually subducted back into the mantle. Continental lithosphere is thicker and lighter than oceanic lithosphere and not easily subducted.

be more correct to think of plate motion as part of mantle convection. Where plates move apart, mantle rises to fill the gap. As the mantle does so, it melts. The melt rises to the surface as magma and creates new oceanic crust at volcanoes along mid-ocean ridges (Figure 1.10). Mid-ocean ridges, such as the East Pacific Rise and the Mid-Atlantic Ridge, thus mark *divergent plate boundaries*. As the oceanic crust moves away from the mid-ocean ridge it cools, along with the mantle immediately below it. This cooling produces a steadily thickening lithosphere. As this lithosphere cools, it contracts and its density increases. Because of this contraction, the depth of the ocean floor increases away from the mid-ocean ridge. When this lithosphere has cooled sufficiently, after 100 million years or so, it becomes denser than the underlying asthenosphere. The lithosphere may then sink back into the mantle in a process known as *subduction*. As the lithosphere sinks, it creates deep ocean trenches, such as the Peru–Chile Trench, or the Marianas Trench. Chains of volcanoes, known as island arcs, invariably occur adjacent to these deep-sea trenches. The volcanism occurs as a result of dehydration of the subducting oceanic crust and lithosphere. Water released from the subducting oceanic crust rises into the overlying mantle, causing it to melt. The island arcs and deep sea trenches are collectively called *sub-*

duction zones. Subduction zones thus mark *convergent plate boundaries*. It is primarily the sinking of old, cold lithosphere that drives the motion of plates. Thus the lithosphere does not merely ride upon convecting mantle – its motion is actually part of mantle convection.

The density of the continental crust is always lower than that of the mantle, regardless of how cold the crust becomes. As a result, it cannot be subducted into the mantle. The Indian–Eurasian plate boundary is a good example of what happens when two continental plates converge. Neither plate readily subducts and the resulting compression has produced, and continues to uplift, the Himalayan Mountains and the Tibetan Plateau. This area of continental crust is not only high – it is also deep. The crust beneath this region extends to depths of as much as 100 km, nearly three times the average crustal thickness. Rocks within this thickened crust will experience increased temperatures and pressures, leading to *metamorphism*, a process in which new minerals form in place of the original ones. In the deepest part of the crust, melting may occur, giving rise to granitic magmas, which will then intrude into the upper crust.

The topographically high Himalayas are subject to extremely high rates of erosion, and the rivers draining the area carry enormous quantities of sediment. These are deposited mainly in the northern Indian Ocean, building the Ganges and Indus Fans outward from the continental margin. As the mountains erode, the mass of crust bearing down on the underlying asthenosphere is reduced. As a result of the decreased downward force, further uplift occurs.

The third kind of plate boundary is known as a transform boundary and occurs where plates slide past one another. A good example of this type of plate boundary is the San Andreas Fault system of California. Here the Pacific Plate is sliding northward past the North American Plate. The passage is not an easy one, however. The two plates occasionally stick together. When they do, stresses steadily build up. Eventually, the stress exceeds the frictional forces holding the plates together, and there is a sudden jump producing an earthquake. Earthquakes are also common in subduction zones and along mid-

ocean ridges. They are much rarer in the interior of plates.

Most volcanism and crustal deformation occur along plate boundaries. A few volcanoes, however, are located in plate interiors and appear to be entirely unrelated to plate tectonic processes. Crustal uplift also occurs in association with these volcanoes. Two good examples are Hawaii and Yellowstone. These phenomena are thought to be the result of mantle plumes. Mantle plumes are convective upwellings. In contrast to the convective upwelling occurring along mid-ocean ridges, which is typically sheet-like, mantle plumes appear to be narrow (~100 km diameter) and approximately cylindrical. Furthermore, it appears that mantle plumes rise from much deeper in the mantle, perhaps even the core–mantle boundary, than convection associated with plate motion.

1.5.3 Earth materials

The most abundant elements in the Earth are O and Fe (both close to 32%), Mg (~15%), Si (~14%), Ni (~1.8%), Ca (1.7%), and Al (1.6%). The majority of the Earth's Fe and Ni are in the core. The remaining rocky part of the Earth, the mantle and crust, consists of ~44% O, ~23% Mg, ~21% Si, ~8% Fe, ~2.5% Ca, and ~2.4% Al. As a consequence, the outer part of the Earth consists principally of compounds known as *silicates*. Silicates are compounds based on the silica tetrahedron, consisting of a silicon atom surrounded by four oxygens (Figure 1.11a). The bonds between the oxygens and silicon are about 50% covalent and 50% ionic and are quite strong. The silicon atom shares an electron with each of the four oxygens. Since oxygen has two valence electrons, each oxygen can form an additional bond. There are two possibilities: the oxygen can form a second bond (which is usually more ionic) with another metal atom or it can form a bond with a second silicon. This latter possibility leads to linking of silica tetrahedra to form rings, chains, sheets, or frameworks. Oxygens bound to two silicons are called *bridging oxygens*.

In orthosilicates, the silica tetrahedra are either completely independent or form dimers, that is, two linked tetrahedra. A good example of a mineral of this type is olivine, whose

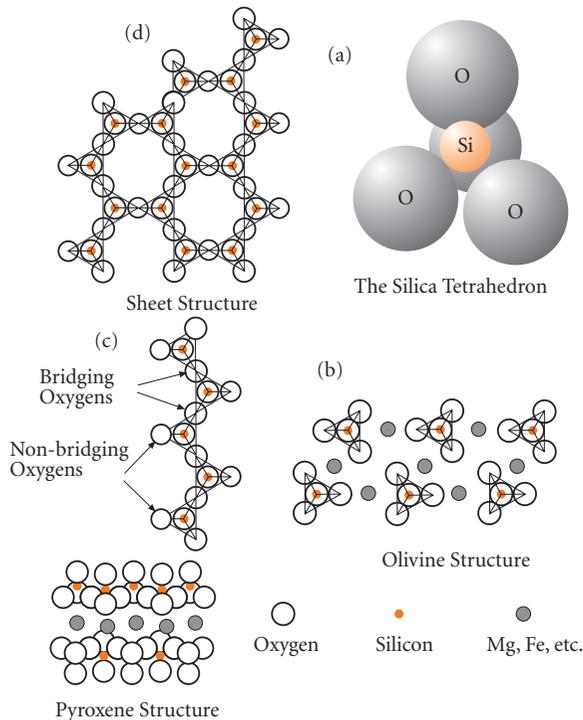


Figure 1.11 The silica tetrahedron and the structure of silicate minerals. (a) The silica tetrahedron consists of a central silicon atom bound to four oxygens. (b) In orthosilicates such as olivine, the tetrahedra are separate and each oxygen is also bound to other metal ions that occupy interstitial sites between the tetrahedra. (c) In pyroxenes, the tetrahedra each share two oxygens and are bound together into chains. Metal ions are located between the chains. (d) In sheet silicates, such as talc, mica, and clays, the tetrahedra each share three oxygens and are bound together into sheets.

structure is illustrated in Figure 1.11b. The chemical formula for olivine is $(\text{Mg,Fe})_2\text{SiO}_4$. The notation (Mg,Fe) indicates that either magnesium or iron may be present. *Olivine* is an example of a solid solution between the Mg end-member, forsterite (Mg_2SiO_4), and the Fe end-member, fayalite (Fe_2SiO_4). Such solid solutions are quite common among silicates. As the formula indicates, there are two magnesium or iron atoms for each silica tetrahedra. Since each Mg or Fe has a charge of +2, their charge balances the -4 charge of each silica tetrahedra. Olivine constitutes roughly 50% of the Earth's upper mantle, and

is thus one of the most abundant minerals in the Earth. At great pressures, it breaks down to a Fe-Mg oxide (magnesiowüstite) and $(\text{Mg,Fe})\text{SiO}_3$ (Mg-perovskite). At the same time, silicon atoms become octahedrally coordinated (surrounded by 6 oxygens rather than 4 as in the silica tetrahedra).

In chain silicates, the silica tetrahedra are linked together to form infinite chains (Figure 1.11c), with two bridging oxygen per tetrahedra. Minerals of this group are known as *pyroxenes* and have the general formula XSiO_3 where X is some metal, usually Ca, Mg, or Fe, which is located between the chains. Two pyroxenes, orthopyroxene ($(\text{Mg,Fe})\text{SiO}_3$) and clinopyroxene ($\text{Ca}(\text{Mg,Fe})\text{Si}_2\text{O}_6$) are very abundant in the Earth's upper mantle as well as in mafic igneous rocks. The pyroxenes wollastonite (CaSiO_3) and jadeite ($\text{NaAlSi}_2\text{O}_6$) are found exclusively in metamorphic rocks.

In double chain silicates, an additional one half bridging oxygen per tetrahedra joins two chains together. Minerals of this group are known as *amphiboles*, which occur widely in both igneous and metamorphic rocks. Among the important minerals in this group are hornblende ($\text{Ca}_2\text{Na}(\text{Mg,Fe})_4\text{Al}_3\text{Si}_8\text{O}_{22}(\text{OH})_2$), tremolite-actinolite ($\text{Ca}_2(\text{Mg,Fe})_5\text{Si}_8\text{O}_{22}(\text{OH})_2$), and glaucophane ($\text{Ca}_2(\text{Mg,Fe})_3\text{Al}_3\text{Si}_8\text{O}_{22}(\text{OH})_2$). These minerals all contain OH as an essential component (Cl or F sometimes substitutes for OH). They are thus examples of hydrous silicates.

Sharing of a third oxygen links the tetrahedra into sheets, forming the sheet silicates (Figure 1.11d). This group includes *micas* such as biotite ($\text{K}(\text{Mg,Fe})_3\text{AlSi}_3\text{O}_{10}(\text{OH})_2$) and muscovite ($\text{KAl}_3\text{Si}_3\text{O}_{10}(\text{OH})_2$), talc ($\text{Mg}_3\text{Si}_4\text{O}_{10}(\text{OH})_2$), and *clay minerals* such as kaolinite ($\text{Al}_2\text{Si}_2\text{O}_5(\text{OH})_4$). As in amphiboles, OH is an essential component of sheet silicates. These minerals can form through weathering and are thus primary sedimentary minerals. Many of them are found in igneous and metamorphic rocks as well.

When all four oxygens are shared between tetrahedra, the result is a framework. The simplest framework silicate is quartz (SiO_2), which consists solely of linked SiO_4 tetrahedra. The other important group of framework silicates is the *feldspars*, of which there are three end-members: sanidine (KAlSi_3O_8), albite ($\text{NaAlSi}_3\text{O}_8$), and anorthite ($\text{CaAl}_2\text{Si}_2\text{O}_8$). The

calcium and sodium feldspars form the plagioclase solid solution, which is stable through a large temperature range. Sodium and potassium feldspars, collectively called alkali feldspar, form more limited solid solutions. Feldspars are the most abundant minerals in the Earth's crust.

Silicates are the most abundant minerals in the Earth, but not the only ones. Other classes of minerals include oxides such as magnetite (Fe_3O_4) and ilmenite (FeTiO_3), carbonates such as calcite (CaCO_3), sulfates such as gypsum ($\text{CaSO}_4 \cdot 2\text{H}_2\text{O}$), hydroxides such as gibbsite ($\text{Al}(\text{OH})_3$), and sulfides such as pyrite FeS_2 .

REFERENCES AND SUGGESTIONS FOR FURTHER READING

- Albarède, F. 2009. *Geochemistry: An Introduction*. New York, Cambridge University Press.
- Broecker, W.S. and Oversby, V.M. 1971. *Chemical Equilibria in the Earth*. New York, McGraw-Hill.
- Gribbin, J.R. 1984. *In Search of Schrödinger's Cat: Quantum Physics and Reality*. New York, Bantam Books.
- Lindley, D. 2001. *Boltzmann's Atom: The Great Debate that Launched a Revolution in Physics*. New York, Free Press.
- Mason, S.F. 1991. *Chemical Evolution*. Oxford, Clarendon Press.
- McSween, H.Y., Richardson, S.H. and Uhle, M. 2003. *Geochemistry: Pathways and Processes*. New York, Columbia University Press.
- Morris, R. 2003. *The Last Sourcers: The Path from Alchemy to the Periodic Table*. Washington, DC, Joseph Henry Press.
- Strathern, P. 2000. *Mendeleev's Dream: The Quest for the Elements*. London, Berkley.
- Walther, J.V. 2009. *Essentials of Geochemistry*. Sudbury, MA, Jones and Bartlett.

Chapter 2

Energy, entropy and fundamental thermodynamic concepts

2.1 THE THERMODYNAMIC PERSPECTIVE

We defined geochemistry as the application of chemical knowledge and techniques to solve geologic problems. It is appropriate, then, to begin our study of geochemistry with a review of physical chemistry. Our initial focus will be on *thermodynamics*. Strictly defined, thermodynamics is the study of energy and its transformations. Chemical reactions and changes of states of matter inevitably involve energy changes. By using thermodynamics to follow the energy, we will find that we can predict the outcome of chemical reactions, and hence the state of matter in the Earth. In principle at least, we can use thermodynamics to predict at what temperature a rock will melt and the composition of that melt, and we can predict the sequence of minerals that will crystallize to form an igneous rock from the melt. We can predict the new minerals that will form when that igneous rock undergoes metamorphism, and we can predict the minerals and the composition of the solution that forms when that metamorphic rocks weathers. Thus thermodynamics allows us to understand (in the sense that we defined understanding in Chapter 1) a great variety of geologic processes.

Thermodynamics embodies a *macroscopic* viewpoint, that is, it concerns itself with the properties of a system, such as temperature, volume, and heat capacity, and it does not concern itself with how these properties are reflected in the internal arrangement of atoms. The *microscopic* viewpoint, which is con-

cerned with transformations on the atomic and subatomic levels, is the realm of *statistical mechanics* and *quantum mechanics*. In our treatment, we will focus mainly on the macroscopic (thermodynamic) viewpoint, but we will occasionally consider the microscopic (statistical mechanical) viewpoint when our understanding can be enhanced by doing so.

In principle, *thermodynamics is only usefully applied to systems at equilibrium*. If an equilibrium system is perturbed, thermodynamics can predict the new equilibrium state, but cannot predict how, how fast, or indeed whether, the equilibrium state will be achieved. (The field of *irreversible thermodynamics*, which we will not treat in this book, attempts to apply thermodynamics to non-equilibrium states. However, we will see in Chapter 5 that thermodynamics, through the *principle of detailed balancing* and *transition state theory*, can help us predict reaction rates.)

Kinetics is the study of rates and mechanisms of reaction. Whereas thermodynamics is concerned with the ultimate equilibrium state and not concerned with the pathway to equilibrium, kinetics concerns itself with the pathway to equilibrium. Very often, equilibrium in the Earth is not achieved, or achieved only very slowly, which naturally limits the usefulness of thermodynamics. Kinetics helps us to understand how equilibrium is achieved and why it is occasionally not achieved. Thus these two fields are closely related, and together form the basis of much of geochemistry. We will treat kinetics in Chapter 5.

2.2 THERMODYNAMIC SYSTEMS AND EQUILIBRIUM

We now need to define a few terms. We begin with the term *system*, which we have already used. A thermodynamic system is simply that part of the universe we are considering. Everything else is referred to as the *surroundings*. A thermodynamic *system* is defined at the convenience of the observer in a manner so that thermodynamics may be applied. While we are free to choose the boundaries of a system, our choice must nevertheless be a careful one as the success or failure of thermodynamics in describing the system will depend on how we have defined its boundaries. Thermodynamics often allows us this sort of freedom of definition. This can certainly be frustrating, particularly for someone exposed to thermodynamics for the first time (and often even the second or third time). But this freedom allows us to apply thermodynamics successfully to a much broader range of problems than otherwise.

A system may be related to its environment in a number of ways. An *isolated* system can exchange neither energy (heat or work) nor matter with its surroundings. A truly isolated system does not exist in nature, so this is strictly a theoretical concept. An *adiabatic* system can exchange energy in the form of work, but not heat or matter, with its surroundings, that is to say it has thermally insulating boundaries. Though a truly adiabatic system is probably also a fiction, heat transport in many geologic systems is sufficiently slow that they may be considered adiabatic. *Closed* systems may exchange energy, in the form of both heat and work with their surrounding but cannot exchange matter. An *open* system may exchange both matter and energy across its boundaries. The various possible relationships of a system to its environment are illustrated in Figure 2.1.

Depending on how they behave over time, systems are said to be either in *transient* or *time-invariant* states. Transient states are those that change with time. Time-independent states may be either *static* or *dynamic*. A dynamic time-independent state, or *steady-state*, is one whose thermodynamic and chemical characteristics do not change with time despite internal changes or exchanges of mass and energy with its surroundings. As we will

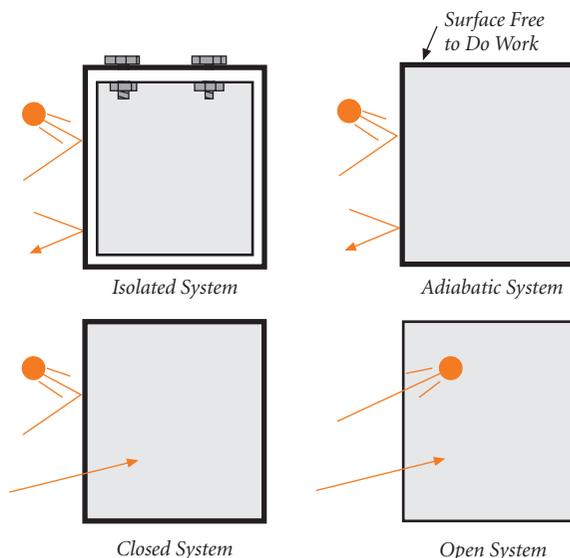


Figure 2.1 Systems in relationship to their surroundings. The ball represents mass exchange, the arrow represents energy exchange.

see, the ocean is a good example of a steady-state system. Despite a constant influx of water and salts from rivers and loss of salts and water to sediments and the atmosphere, its composition does not change with time (at least on geologically short time-scales). Thus a steady-state system may also be an open system. We could define a static system as one in which nothing is happening. For example, an igneous rock or a flask of seawater (or some other solution) is static in the macroscopic perspective. From the statistical mechanical viewpoint, however, there is a constant reshuffling of atoms and electrons, but with no net changes. Thus static states are generally also dynamic states when viewed on a sufficiently fine scale.

Let's now consider one of the most important concepts in physical chemistry, that of *equilibrium*. One of the characteristics of the *equilibrium state* is that it is static from a macroscopic perspective, that is, it does not change measurably with time. Thus the equilibrium state is always time-invariant. However, while a reaction $A \rightarrow B$ may appear to have reached static equilibrium on a macroscopic scale, this reaction may still proceed on a microscopic scale but with the rate of reaction $A \rightarrow B$ being the same as that of $B \rightarrow A$. Indeed, a kinetic definition of equilibrium is

that the forward and reverse rates of reaction are equal.

The equilibrium state is entirely independent of the manner or pathway in which equilibrium is achieved. Indeed, once equilibrium is achieved, no information about previous states of the system can be recovered from its thermodynamic properties. Thus a flask of CO_2 produced by combustion of graphite cannot be distinguished from CO_2 produced by combustion of diamond. In achieving a new equilibrium state, all records of past states are destroyed.

Time-invariance is a necessary but not sufficient condition for equilibrium. Many systems exist in metastable states. Diamond at the surface of the Earth is not in an equilibrium state, despite its time-invariance on geologic time-scales. Carbon exists in this metastable state because of kinetic barriers that inhibit transformation to graphite, the equilibrium state of pure carbon at the Earth's surface. Overcoming these kinetic barriers generally requires energy. If diamond is heated sufficiently, it will transform to graphite, or in the presence of sufficient oxygen, to CO_2 .

The concept of equilibrium versus metastable or unstable (transient) states is illustrated in Figure 2.2 by a ball on a hill. The equilibrium state is when the ball is in the valley at the bottom of the hill, because its gravitational potential energy is minimized in this position. When the ball is on a slope, it is in an unstable, or transient, state and will tend to roll down the hill. However, it may also become trapped in small depressions on the side of the hill, which represent metastable states. The small hill bordering the depression represents a kinetic barrier. This kinetic barrier

can only be overcome when the ball acquires enough energy to roll up and over it. Lacking that energy, it will exist in the metastable state indefinitely.

In Figure 2.2, the ball is at equilibrium when its (gravitational) potential energy is lowest (i.e., at the bottom of the hill). This is a good definition of equilibrium in this system, but as we will soon see, it is not adequate in all cases. A more general statement would be to say that *the equilibrium state is the one toward which a system will change in the absence of constraints*. So in this case, if we plane down the bump (remove a constraint), the ball rolls to the bottom of the hill. At the end of this Chapter, we will be able to produce a thermodynamic definition of equilibrium based on the *Gibbs free energy*. We will find that, for a given pressure and temperature, the chemical equilibrium state occurs when the Gibbs free energy of the system is lowest.

Natural processes proceeding at a finite rate are *irreversible* under a given set of conditions; that is, they will only proceed in one direction. Here we encounter a problem in the application of thermodynamics: if a reaction is proceeding, then the system is out of equilibrium and thermodynamic analysis cannot be applied. This is one of the first of many paradoxes in thermodynamics. This limitation might at first seem fatal, but we get around it by imagining a comparable reversible reaction. *Reversibility* and *local equilibrium* are concepts that allow us to “cheat” and apply thermodynamics to non-equilibrium situations. A “reversible” process is an idealized one where the reaction proceeds in sufficiently small steps that it is in equilibrium at any given time (thus allowing the application of thermodynamics).

Local equilibrium embodies the concept that in a closed or open system, which may not be at equilibrium on the whole, small volumes of the system may none the less be at equilibrium. There are many such examples. In the example of a mineral crystallizing from magma, only the rim of the crystal may be in equilibrium with the melt. Information about the system may nevertheless be derived from the relationship of this rim to the surrounding magma. Local equilibrium is in a sense the spatial equivalent to the temporal concept of reversibility and allows the application of thermodynamics to real systems,

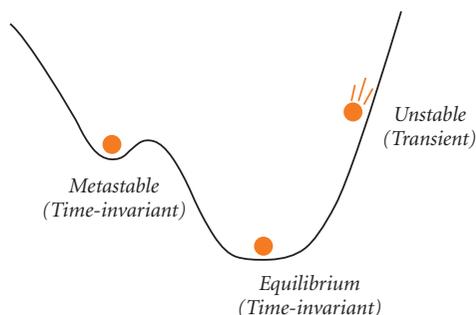


Figure 2.2 States of a system.

which are invariably non-equilibrium at large scales. Both local equilibrium and reversibility are examples of simplifying assumptions that allow us to treat complex situations. In making such assumptions, some accuracy in the answer may be lost. Knowing when and how to simplify a problem is an important scientific skill.

2.2.1 Fundamental thermodynamic variables

In the next two chapters we will be using a number of variables, or properties, to describe thermodynamic systems. Some of these will be quite familiar to you, others less so. Volume, pressure, energy, heat, work, entropy, and temperature are the most fundamental variables in thermodynamics. As all other thermodynamic variables are derived from them, it is worth our while to consider a few of these properties.

Energy is the capacity to produce change. It is a fundamental property of any system, and it should be familiar from physics. By choosing a suitable reference frame, we can define an absolute energy scale. However, it is changes in energy that are generally of interest to us rather than absolute amounts. Work and heat are two of many forms of energy. Heat, or thermal energy, results from random motions of molecules or atoms in a substance and is closely related to kinetic energy. Work is done by moving a mass, M , through some distance, $x = X$, *against* a force F :

$$w = -\int_0^x F dx \quad (2.1)$$

where w is work and force is defined as mass times acceleration:

$$F = -M \frac{dv}{dt} \quad (2.2)$$

(the minus signs are there because of the convention that *work done on a system is positive, work done by a system is negative*). This is, of course, Newton's first law. In chemical thermodynamics, pressure–volume work is usually of more interest. Pressure is defined as force per unit area:

$$P = \frac{F}{A} \quad (2.3)$$

Since volume is area times distance, we can substitute equation 2.3 and $dV = A dx$ into eqn. 2.1 and obtain:

$$w = -\int_{x_0}^{x_1} \frac{F}{A} A dx = -\int_{V_0}^{V_1} P dV \quad (2.4)$$

Thus work is also done as a result of a volume change in the presence of pressure.

Potential energy is energy possessed by a body by virtue of its position in a force field, such as the gravitational field of the Earth, or an electric field. Chemical energy will be of most interest to us in this book. Chemical energy is a form of potential energy stored in chemical bonds of a substance. Chemical energy arises from the electromagnetic forces acting on atoms and electrons. Internal energy, which we denote with the symbol U , is the sum of the potential energy arising from these forces as well as the kinetic energy of the atoms and molecules (i.e., thermal energy) in a substance. It is internal energy that will be of most interest to us.

We will discuss all these fundamental variables in more detail in the next few sections.

2.2.1.1 Properties of state

Properties or variables of a system that depend only on the present state of the system, and not on the manner in which that state was achieved, are called *variables of state* or *state functions*. *Extensive* properties depend on total size of the system. Mass, volume, and energy are all extensive properties. Extensive properties are additive, the value for the whole being the sum of values for the parts. *Intensive* properties are independent of the size of a system, for example temperature, pressure, and viscosity. They are not additive; for example, the temperature of a system is not the sum of the temperature of its parts. In general, an extensive property can be converted to an intensive one by dividing it by some other extensive property. For example, density is the mass per volume and is an intensive property. It is generally more convenient to work with intensive rather than extensive properties. For a single component system not undergoing reaction, specification of three variables (two intensive, one extensive) is generally sufficient to determine the rest, and specification of any two intensive variables is

generally sufficient to determine the remaining intensive variables.

A final definition is that of a *pure substance*. A pure substance is one that cannot be separated into fractions of different properties by the same processes as those considered. For example, in most processes, the compound H₂O can be considered a pure substance. However, if electrolysis were involved, this would not be the case.

2.3 EQUATIONS OF STATE

Equations of state describe the relationship that exists among the state variables of a system. We will begin by considering the ideal gas law and then very briefly consider two more complex equations of state for gases.

2.3.1 Ideal gas law

The simplest and most fundamental of the equations of state is the *ideal gas law*.^{*} It states that pressure, volume, temperature, and the number of moles of a gas are related as:

$$\boxed{PV = NRT} \quad (2.5)$$

where P is pressure, V is volume, N is the number of moles, T is thermodynamic, or absolute temperature (which we will explain shortly), and R is the ideal gas constant[†] (an empirically determined constant equal to 8.314 J/mol-K, 1.987 cal/mol-K or 82.06 cc-atm/deg-mol). This equation describes the relation between two extensive (mass-dependent) parameters, volume and the number of moles, and two intensive (mass-independent) parameters, temperature and pressure. We earlier stated that if we defined two intensive and one extensive system parameter, we could determine the remaining parameters. We can see from eqn. 2.5 that this is indeed the case for an ideal gas. For example, if we know N , P , and T , we can use eqn. 2.5 to determine V .

The ideal gas law, and any equation of state, can be rewritten with intensive properties only. Dividing V by N we obtain the

molar volume, \bar{V} . Substituting \bar{V} for V and rearranging, the ideal gas equation becomes:

$$\bar{V} = \frac{RT}{P} \quad (2.6)$$

The ideal gas equation tells us how the volume of a given amount of gas will vary with pressure and temperature. To see how molar volume will vary with temperature alone, we can differentiate eqn. 2.6 with respect to temperature, holding pressure constant, and obtain:

$$\left(\frac{\partial V}{\partial T}\right)_P = \frac{\partial(NRT/P)}{\partial T} \quad (2.7)$$

which reduces to:

$$\left(\frac{\partial V}{\partial T}\right)_P = \frac{NR}{P} \quad (2.8)$$

It would be more useful to know the *fractional* volume change rather than the absolute volume change with temperature, because the result in that case does not depend on the size of the system. To convert to the fractional volume change, we simply divide the equation by V :

$$\frac{1}{V} \left(\frac{\partial V}{\partial T}\right)_P = \frac{NR}{PV} \quad (2.9)$$

Comparing eqn. 2.9 with 2.5, we see that the right-hand side of the equation is simply $1/T$, thus

$$\frac{1}{V} \left(\frac{\partial V}{\partial T}\right)_P = \frac{1}{T} \quad (2.10)$$

The left-hand side of this equation, the fractional change in volume with change in temperature, is known as the *coefficient of thermal expansion*, α :

$$\alpha \equiv \frac{1}{V} \left(\frac{\partial V}{\partial T}\right)_P \quad (2.11)$$

^{*} Frenchman Joseph Gay-Lussac (1778–1850) established this law based on the earlier work of Englishman Robert Boyle and Frenchman Edme Mariotte.

[†]We will generally refer to it merely as the gas constant.

For an ideal gas, the coefficient of thermal expansion is simply the inverse of temperature.

The *compressibility* of a substance is defined in a similar manner as the fractional change in volume produced by a change in pressure at constant temperature:

$$\beta \equiv -\frac{1}{V} \left(\frac{\partial V}{\partial P} \right)_T \quad (2.12)$$

Geologists sometimes use the *isothermal bulk modulus*, K_T , in place of compressibility. The isothermal bulk modulus is simply the inverse of compressibility: $K_T = 1/\beta$. Through a similar derivation to the one we have just done for the coefficient of thermal expansion, it can be shown that the compressibility of an ideal gas is $\beta = 1/P$.

The ideal gas law can be derived from statistical physics (first principles), *assuming the molecules occupy no volume and have no electrostatic interaction*. Doing so, we find, $R = N_0 k$, where k is *Boltzmann's constant* (1.381×10^{-23} J/K) and N_0 is Avagadro's Number (the number of atoms in one mole of a substance). k is a fundamental constant that relates the average molecular energy, e , of an ideal gas to its temperature (K) as $e = 3kT/2$.

Since the assumptions just stated are ultimately invalid, it is not surprising that the ideal gas law is only an approximation for real gases; it applies best in the limit of high temperature and low pressure. Deviations are largest near the condensation point of the gas.

The *compressibility factor* is a measure of deviation from ideality and is defined as

$$Z = PV/NRT \quad (2.13)$$

By definition, $Z = 1$ for an ideal gas.

2.3.2 Equations of state for real gases

2.3.2.1 Van der Waals equation

Factors we need to consider in constructing an equation of state for a real gas are the finite volume of molecules and the attractive and repulsive forces between molecules arising from electric charges. The Van der Waals equation is probably the simplest equation of state that takes account of these factors. The Van der Waals equation is:

$$P = \frac{RT}{\bar{V} - b} - \frac{a}{\bar{V}^2} \quad (2.14)$$

Here again we have converted volume from an extensive to an intensive property by dividing by N .

Let's examine the way in which the Van der Waals equation attempts to take account of finite molecular volume and forces between molecules. Considering first the forces between molecules, imagine two volume elements v_1 and v_2 . The attractive forces will be proportional to the number of molecules or the concentrations, c_1 and c_2 , in each. Therefore, attractive forces are proportional to $c_1 \times c_2 = c^2$. Since c is the number of molecules per unit volume, $c = n/V$, we see that attractive forces are proportional to $1/\bar{V}^2$. Thus it is the second term on the right that takes account of forces between molecules. The a term is a constant that depends on the nature and strength of the forces between molecules, and will therefore be different for each type of gas.

In the first term on the right, \bar{V} has been replaced by $\bar{V} - b$. b is the volume actually occupied by molecules, and the term $\bar{V} - b$ is the volume available for movement of the molecules. Since different gases have molecules of differing size, we can expect that the value of b will also depend on the nature of the gas. Table 2.1 lists the values of a and b for a few common gases.

2.3.2.2 Other equations of state for gases

The *Redlich-Kwong Equation* (1949) expresses the attractive forces as a more complex function:

Table 2.1 Van der Waals constants for selected gases.

Gas	a	b
	liter ² -atm/mole ²	liter/mole
Helium	0.034	0.0237
Argon	1.345	0.0171
Hydrogen	0.244	0.0266
Oxygen	1.360	0.0318
Nitrogen	1.390	0.0391
Carbon dioxide	3.592	0.0399
Water	5.464	0.0305
Benzene	18.00	0.1154

$$P = \frac{RT}{\bar{V} - b} - \frac{a}{T^{1/2}\bar{V}(\bar{V} + b)} \quad (2.15)$$

The *Virial Equation* is much easier to handle algebraically than the van der Waals equation and has some theoretical basis in statistical mechanics:

$$P\bar{V} = A + BP + CP^2 + DP^3 + \dots \quad (2.16)$$

A, B, C, \dots are empirically determined (temperature-dependent) constants.

2.3.3 Equation of state for other substances

The compressibility and coefficient of thermal expansion parameters allow us to construct an equation of state for any substance. Such an equation relates the fundamental properties of the substance: its temperature, pressure, and volume. The partial differential of volume with respect to temperature and pressure is such an equation:

$$dV = \left(\frac{\partial V}{\partial T}\right)_P dT + \left(\frac{\partial V}{\partial P}\right)_T dP \quad (2.17)$$

Substituting the coefficient of thermal expansion and compressibility for $\partial V/\partial T$ and $\partial V/\partial P$ respectively we have:

$$dV = V(\alpha dT - \beta dP) \quad (2.18)$$

Thus to write an equation of state for a substance, our task becomes to determine its compressibility and coefficient of thermal expansion. Once we know them, we can integrate eqn. 2.18 to obtain the equation of state. These, however, will generally be complex functions of temperature and pressure, so the task is often not easy.

2.4 TEMPERATURE, ABSOLUTE ZERO, AND THE ZEROTH LAW OF THERMODYNAMICS

How do you define and measure temperature? We discussed temperature with respect to the ideal gas law without defining it, though we all have an intuitive sense of what temperature is. We noted earlier that temperature of a gas is a measure of the average (kinetic) energy of its molecules. Another approach might be to use the ideal gas law to construct

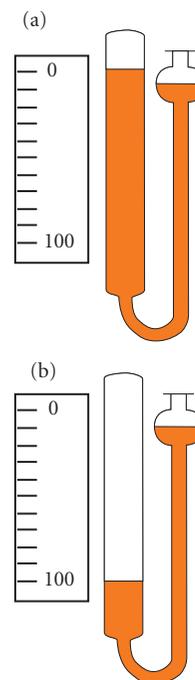


Figure 2.3 An ideal gas thermometer.

a *thermometer* and define a temperature scale. A convenient thermometer might be one based on the linear relationship between temperature and the volume of an ideal gas. Such a thermometer is illustrated in Figure 2.3. The equation describing the relationship between the volume of the gas in the thermometer and our temperature, τ , is:

$$V = V_0(1 + \gamma\tau) \quad (2.19)$$

where V_0 is the volume at some reference point where $\tau = 0$ (Figure 2.3a) and γ is a scale factor. For example, we might choose $\tau = 0$ to be the freezing point of water and the scale factor such that $\tau = 100$ occurs at the boiling point of water (Figure 2.3b), as is the case in the centigrade scale. Rearranging, we have:

$$\tau = \frac{1}{\gamma} \left(\frac{V}{V_0} - 1 \right) \quad (2.20)$$

Then $\tau = 0$ at $V = V_0$. If V is less than the reference volume, then temperature will be negative on our scale. But notice that while any positive value of temperature is possible on this scale, there is a limit to the range of possible negative values. This is because V can

never be negative. The minimum value of temperature on this scale will occur when V is 0. This occurs at:

$$\tau_0 = -\frac{1}{\gamma} \quad (2.21)$$

Thus implicit in the ideal gas law, which we used to make this thermometer, is the idea that there is an absolute minimum value, or an absolute zero, of temperature, which occurs when the volume of an ideal gas is 0. Notice that while the value $(-1/\gamma)$ of this absolute zero will depend on how we designed our thermometer (i.e., on V_0), the result, that a minimum value exists, does not. We should also point out that only an ideal gas can have a volume of 0. The molecules of real gases have a finite volume, and such a gas will have a finite volume at absolute zero.

The temperature scale used by convention in thermodynamics is the Kelvin* scale. The magnitude of units, called kelvins (not degrees kelvin) and designated K (not °K), on this scale are the same as the centigrade scale, so there are exactly 100 kelvins between the freezing and boiling point of water. There is some slight uncertainty (a very much smaller uncertainty than we need to concern ourselves with) concerning the value of absolute zero (i.e., the value of γ in eqns 2.20 and 2.21). The scale has been fixed by choosing 273.16 kelvins to be the triple point of water (0.01°C). On this scale, the absolute zero of temperature occurs at 0 ± 0.01 kelvins. *The Kelvin scale should be used wherever temperature occurs in a thermodynamic equation.*

Temperature has another fundamental property, and this is embodied in the *zerOTH*

law of thermodynamics. It is sufficiently obvious from everyday experience that we might overlook it. It concerns thermal equilibrium and may be stated in several ways: *two bodies in thermal equilibrium have the same temperature and any two bodies in thermal equilibrium with a third are in equilibrium with each other.*

2.5 ENERGY AND THE FIRST LAW OF THERMODYNAMICS

2.5.1 Energy

The first law may be stated in various ways:

- *heat and work are equivalent*
- *energy is conserved in any transformation*
- *the change of energy of a system is independent of the path taken.*[†]

All are restatements of the law of conservation of energy:

Energy can be neither created nor destroyed.

Mathematically:

$$\Delta U = Q + W \text{ or } dU = dQ + dW \quad (2.22)$$

Thermodynamics is concerned only with the internal energy of a system. We don't really care whether the system as a whole is in motion, i.e., whether it has kinetic energy (we do care, however, about the internal kinetic energy, or heat). For the most part, we also don't care whether it has potential energy, except to the extent that this influences the

* Named for Lord Kelvin. Born William Thomson in Scotland in 1824, he was appointed Professor at Glasgow University at the age of 22. Among his many contributions to physics and thermodynamics was the concept of absolute temperature. He died in 1907.

[†]This may seem intuitively obvious to us, but it was not to James Joule (1818–1889), English brewer and physicist, who postulated it on the basis of experimental results. It was not obvious to his contemporaries either. His presentation of the idea of equivalence of heat and work to the British Association in 1843 was received with “entire incredulity” and “general silence”. The Royal Society rejected his paper on the subject a year later. If you think about it a bit, it is not so obvious – in fact, there is no good reason why heat and work should be equivalent. This law is simply an empirical observation. The proof is a negative one: experience has found no contradiction of it. German physician Julius Mayer (1814–1878) formulated the idea of conservation of energy in 1842, but his writing attracted little attention. It was Joule's experiments with heat and work that conclusively established the principle of conservation of energy. By 1850, the idea of conservation of energy began to take hold among physicists, thanks to Joule's persistence and the support of a brilliant young physicist named William Thomson, who also had been initially skeptical.

state of our system (e.g., pressure in the atmosphere is a function of the altitude, and hence would be of interest to us). In addition, we are almost always concerned *only* with energy *changes*, not with the absolute energy of a system. In thermodynamics, ΔU , not U , is the interesting quantity.

Energy may be transferred between a system and its surroundings in several ways: heat, work, radiation, and advection (i.e., energy associated with mass gained or lost by the system). Whenever possible, we will want to choose our system such that it is closed and we don't have to worry about the latter. In most, but not all, instances of geochemical interest, radiation is not important. Thus in geochemical thermodynamics, *heat and work are the forms of energy flow of primary interest.*

2.5.2 Work

We have seen that work is the integral of force applied through a distance. Force times distance has units of energy (mass-velocity²), thus work is a form of energy. The SI (*Systeme Internationale*) unit of energy is the joule = 1 kg-m²/s². Conversion factors for energy and other variables as well as values of important constants are listed in Appendix I.

There are several kinds of work of interest to thermodynamics, the most important of which is that involved in chemical reactions (later, when we consider oxidation and reduction reactions, we will be concerned with electrochemical work). One of the most important forms of work in classical thermodynamics is "PV" work: expansion and contraction. Expressing eqn. 2.4 in differential form:

$$dW = -P_{ext}dV \quad (2.23)$$

Pressure is force per unit area, and therefore has units of mass-time⁻²-distance⁻¹, while volume has units of distance³. The product of P and V therefore has units of energy: mass-(distance/time)².* The negative sign

arises because, *by convention, we define energy flowing into the system as positive.* Work done by the system is thus negative, while work done on the system is positive. This conforms to a 1970 IUPAC (International Union of Pure and Applied Chemistry) recommendation.

While "PV" work is not as important in geochemistry as in other applications of thermodynamics, it is nevertheless of significant interest. There is, of course, a great range of pressures within the Earth. Systems rising within the Earth, such as magma, a hydrothermal fluid, or upwelling water in the ocean, will thus do work on their surroundings, and systems sinking, such as sediments being buried or lithosphere being subducted, will have work done on them.

We mentioned the concept of reversible and irreversible reactions, and stated that a reversible reaction is one that occurs in sufficiently small steps that equilibrium is maintained. In an expansion or contraction reaction, equilibrium is maintained and the reaction is reversible if the external pressure is equal to the internal pressure. The work done under these conditions is said to be reversible:

$$dW_{rev} = -PdV \quad (2.24)$$

2.5.3 Path independence, exact differentials, state functions, and the first law

We said earlier that state functions are those that depend only on the present state of a system. Another way of expressing this is to say that state functions are path-independent. Indeed, path-independence may be used as a test of whether a variable is a state function or not. This is to say that if Y is a state function, then for any process that results in a change $Y1 \rightarrow Y2$, the net change in Y , ΔY , is independent of how one gets from $Y1$ to $Y2$. Furthermore, if Y is a state function, then the differential dY is said to be mathematically *exact*.

* The pascal, the SI unit of pressure, is equal to 1 kg/ms². Thus if pressure is measured in MPa (megapascals, 1 atm \approx 1 bar = 0.1 MPa) and volume in cc (=10⁻⁶ m³), the product of pressure times volume will be in joules. This is rather convenient. It is named for French mathematician and physicist Blaise Pascal (1623–1662). Among his many contributions was the demonstration that atmospheric pressure was lower atop the Puy de Dome volcano than in the town of Clermont-Ferrand below it.

Let's explore what is meant by an exact differential. An exact differential is the familiar kind, the kind we would obtain by differentiating the function u with respect to x and y , and also the kind we can integrate. But not all differential equations are exact. Let's first consider the mathematical definition of an exact differential, then consider some thermodynamic examples of exact and inexact differentials.

Consider the first order differential expression:

$$Mdx + Ndy \quad (2.25)$$

containing variables M and N , which may or may not be functions of x and y . Equation 2.25 is said to be an *exact differential* if there exists some function u of x and y relating them such that the expression:

$$du = Mdx + Ndy \quad (2.26)$$

is the total differential of u :

$$du = \left(\frac{\partial u}{\partial x}\right)_y dx + \left(\frac{\partial u}{\partial y}\right)_x dy \quad (2.27)$$

Let's consider what this implies. Comparing 2.26 and 2.27, we see that:

$$\frac{\partial u}{\partial x} = M \quad \text{and} \quad \frac{\partial u}{\partial y} = N \quad (2.28)$$

A necessary, but not sufficient, condition for 2.25 to be an exact differential is that M and N must be functions of x and y .

A general property of partial differentials is the *reciprocity relation* or *cross-differentiation identity*, which states that the order of differentiation does not matter, so that:

$$\frac{\partial^2 u}{\partial x \partial y} = \frac{\partial^2 u}{\partial y \partial x} \quad (2.29)$$

(The reciprocity relation is an important and useful property in thermodynamics, as we shall see at the end of this chapter.) If eqn. 2.26 is the total differential of u , it follows that:

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x} \quad (2.30)$$

which is equivalent to:

$$\left(\frac{dM}{dy}\right)_x = \left(\frac{dN}{dx}\right)_y \quad (2.31)$$

Equation 2.31 is a necessary and sufficient condition for 2.25 to be an exact differential; that is, if the cross-differentials are equal, then the differential expression is exact.

Exact differentials have the property that they can be integrated and an exact value obtained. This is true because they depend only on the initial and final values of the independent variables (e.g., x and y in eqn. 2.27).

Now let's consider some thermodynamic examples. Volume is a state function and we can express it as an exact differential in terms of other state functions, as in eqn. 2.17:

$$dV = \left(\frac{\partial V}{\partial T}\right)_P dT + \left(\frac{\partial V}{\partial P}\right)_T dP \quad (2.17)$$

Substituting the coefficient of thermal expansion and compressibility for $\partial V/\partial T$ and $\partial V/\partial P$ respectively, equation 2.30 becomes equal to eqn. 2.18:

$$dV = \alpha V dT - \beta V dP \quad (2.18)$$

According to eqn. 2.31, if V is a state function, then:

$$\frac{\partial(\alpha V)}{\partial P} = -\frac{\partial(\beta V)}{\partial T} \quad (2.32)$$

You should satisfy yourself that eqn. 2.32 indeed holds for ideal gases and therefore that V is a state variable.

Work is not a state function, that is, the work done does not depend only on the initial and final states of a system. We would expect then that dW is not an exact differential, and indeed, this is easily shown for an ideal gas.

For PV work, $dW = -PdV$. Substituting eqn. 2.17 for dV and rearranging, we have:

$$dW = -P \left[\left(\frac{\partial V}{\partial T}\right)_P dT + \left(\frac{\partial V}{\partial P}\right)_T dP \right] \quad (2.33)$$

Evaluating $\partial V/\partial T$ and $\partial V/\partial P$ for the ideal gas equation and multiplying through by P , this becomes:

$$dW = -NRdT + \frac{NRT}{P} dP \quad (2.34)$$

but

$$\frac{\partial NR}{\partial P} \neq \frac{\partial(NRT/P)}{\partial T} \quad (2.35)$$

We cannot integrate eqn. 2.34 and obtain a value for the work done without additional knowledge of the variation of T and P because the amount of work done does not depend only on the initial and final values of T and P ; it depends on the path taken. Heat is also not a state function, not an exact differential, and is also path-dependent. Path-dependent functions always have inexact differentials; path-independent functions always have exact differentials.

On a less mathematical level, let's consider how the work and heat will vary in a transformation of a system, say from state 1 to state 2. Imagine that we burn gasoline in an open container. In this case, in the transformation from state 1 (gasoline) to state 2 (combustion products of gasoline), energy is given up by the system only as heat. Alternatively, we could burn the gasoline in an engine and recover some of the energy as work (expansion of the volume of the cylinder resulting in motion of the piston). The end states of these two transformations are the same, but the amount of heat released and work done varied depending on the path we took. Thus neither work nor heat can be state functions. Energy is a state function, is path-independent, and is an exact differential. Whether we burn the gasoline in an open container or an engine, the energy released will be the same. Herein lies the significance for thermodynamics of Joule's discovery: that *the sum of heat and work is independent of the path taken even though, independently, work and heat are not.*

* Rudolf Clausius (1822–1888), a physicist at the Prussian military engineering academy in Berlin, formulated what we now refer to as the second law and the concept of entropy in a paper published in 1850. Similar ideas were published a year later by William Thomson (Lord Kelvin), who is responsible for the word “entropy”. Clausius was a theorist who deserves much of the credit for founding what we now call “thermodynamics” (he was responsible for, among many other things, the virial equation for gases). However, a case can be made that Sadi Carnot (1796–1832) should be given the credit. Carnot was a Parisian military officer (the son of a general in the French revolutionary army) interested in the efficiency of steam engines. The question of credit hinges on whether he was referring to what we now call entropy when he used the word “calorique”.

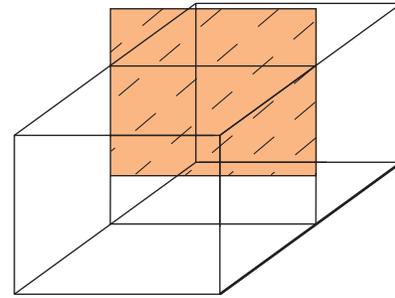


Figure 2.4 A gas-filled box with a removable partition. When the partition is removed, the gases mix. Entropy increases during this process.

2.6 THE SECOND LAW AND ENTROPY

2.6.1 Statement

Imagine a well-insulated box (an isolated system) somewhere in the universe (Figure 2.4). Imagine that within the box are two gases, separated by a removable partition. If we remove the partition, what happens? You know: the two gases mix completely. The process is entirely spontaneous. We have neither added energy to nor taken energy from the system, hence the first law says nothing about this process. Nor did removing the partition “cause” the reaction. This is apparent from the observation that if we reinsert the partition, the gases do not unmix. That you knew that the gases would mix (and knew as well that they would not unmix upon reinserting the partition) suggests there is something very fundamental and universal about this. We need a physical law that describes it. This is the second law.

The second law may be stated in a number of ways:

*It is impossible to construct a machine that is able to convey heat by a cyclical process from one reservoir at a lower temperature to another at a higher temperature unless work is done by some outside agency (i.e., air conditioning is never free).**