Chapter 4

ASPECTS OF MATERIALS SCIENCE

This chapter provides a brief description of materials concepts that may be useful in understanding electronic materials. The review is not exhaustive but is intended to provide a minimum (and rather basic) level of familiarity with important concepts used in other chapters. As elsewhere, the reader is referred to the recommended readings for additional background and details.

4.1 STRUCTURES OF MATERIALS

One of the most fundamental issues in understanding a material is the arrangement of the atoms that make it up. Gases consist of widely spaced atoms distributed randomly in a volume. Because the atoms are spaced, on average, beyond the range of normal electronic interactions, their arrangement is random. Furthermore, for a material to be a gas the bonding interactions that would draw atoms together must be smaller than the thermal interactions that drive them apart, therefore bonds among atoms in a gas must be relatively weak. In a liquid or amorphous solid the atoms are nearly as closely packed as they would be in a crystalline solid (hence the volume change upon freezing is modest) but their interatomic angles are rather random and thus there is no regular long-range order. As we saw in Chapter 2, regular crystal structures are essential determining the properties of most semiconductors, although amorphous semiconductors also exist.

Amorphous solids are much like very viscous liquids. Their organization may be relatively good within the first ring of nearest neighbors relative to a given atom.
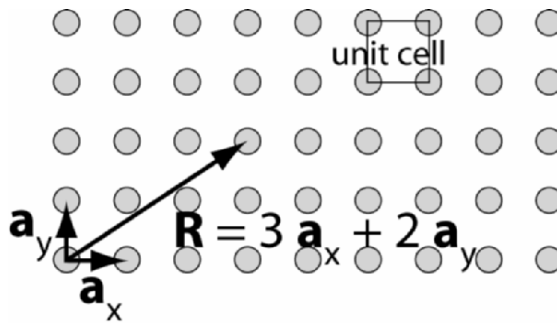
Figure 4.1: A schematic diagram of the periodic arrangement of atoms in a lattice. The basis vectors of the lattice are a_x and a_y. Any lattice point can be written as a linear combination of the two basis vectors.

However, the distance and direction to atoms beyond the first nearest neighbor becomes increasingly uncertain to the point that beyond roughly second-nearest neighbors, the chance of finding an atom in any direction is roughly equal. The simplest way to picture an amorphous solid is to imagine small shapes (triangular pyramids for example) connected through flexible links. The individual shape would be relatively well defined, but the distance and direction from one shape to another would be ill defined and some of the pyramids would be significantly distorted. Many common amorphous solids are based on such a network. Examples of amorphous materials include the semiconductor amorphous silicon and the insulator amorphous silicon dioxide. Amorphous silicon and other amorphous semiconductors are described in more detail in Chapter 8.

4.1.1 Crystal lattices

Crystalline solids form in one of the fourteen basic structural arrangements, known as Bravais lattices. All lattices have one basic property in common – there are a fixed set of translation vectors which will take an observer from any given lattice site to any other site. As we will see below, it is also possible for a lattice site to include more than one atom. Any combination of the basic translation vectors will always end at a lattice site, no matter how far one is from the original position, as long as the crystal contains no defects. Thus, if the translation vectors of a given lattice are $\mathbf{a_1}$, $\mathbf{a_2}$, and $\mathbf{a_3}$, and a reference site is located at coordinate 0,0,0, then all other sites are located at $\mathbf{R} = u\mathbf{a_1} + v\mathbf{a_2} + w\mathbf{a_3}$, where u, v, and w are integers. For example, a square planar lattice, such as the one shown in Figure 4.1, has two translation vectors along the x and y axes. This regularity is important to the collective vibrational motion of atoms in the solid (the phonon modes) and to the nature of electronic states, as we saw in Chapter 2.

In discussing the various directions and planes of atoms in a crystal, it is useful to have a convenient mechanism for naming them. "The planes that run from corner to corner to corner of the cube in a cubic material," would hardly be a helpful name and is imprecise in addition. The labels u, v, and w, by which atom translation vectors were described above, provide a handy index for a direction because any combination of the three defines a direction (as in vector **R**, above) relative to the origin. To indicate a negative direction it is conventional to place a bar over the negative index, rather than using a negative number. This shortens the indexing and keeps the numbers better aligned. Thus, in the index $[1\,\bar{1}\,0]$, the $\bar{1}$ indicates "negative one in the second index direction."

To index planes one may take advantage of the fact that any plane has a single direction perpendicular to it. Therefore the same indexing scheme used for directions can index planes as long as we have a notation by which we may know if we are talking about a plane or a direction. Thus, [u, v, w] with square brackets indicates a direction while we use parentheses, as in (h, k, l), for the plane. If the plane and direction are to be perpendicular it is required that uh+vk+wl=0 in a cubic system. Typical examples of plane indices are shown in Figure 4.2. A brief examination of the cubic lattice in Figure 4.2 shows that there are groups of similar planes. For example, the (100), (010), and (001) planes are all identical except for rotation of the coordinate system as are the corresponding directions. To indicate any of a set of otherwise identical planes one uses braces – thus {100}, while for a set of identical directions, triangular brackets are used as in <100>. One may also note that the (h, k, l) plane intercepts the x, y, and z axes at 1/h, 1/k, and 1/l within the cube for a cubic system. These indices are known as Miller indices.

The fourteen Bravais lattices include three cubic forms (simple, face-centered or body centered), two tetragonal forms (simple or body-centered), four orthorhombics (simple, body-, base-, or face-centered), rhombohedral, hexagonal, two monoclinic (simple or base-centered), and triclinic. All crystal structures, no matter how complex, can be reduced to one of these basic lattices. The most important structures for semiconductors are face-centered cubic (fcc) and hexagonal, although others such as tetragonal and simple cubic are also common. For drawings of these lattices, see Cullity in the recommended readings or a similar crystallography book.

The majority of simple metals form hexagonal, body-centered or face-centered cubic lattices, although other lattices also occur. Which lattice is formed depends upon the details of the atomic orbitals that give rise to bonding in the solid (see Chapter 5). However, there are few obvious trends when one looks at the distribution of lattices on a periodic table. Group IA metals and those with partially filled d-orbitals often form bcc structures, while the noble metals tend to be fcc, as are the group IB metals. Rare earths and metals with slightly filled d-orbitals are primarily hexagonal. Four elements stand out as somewhat unique. C, Si, and Ge as well as the semiconducting form of Sn can all occur in the diamond-structure, a face-centered cubic structure but with two atoms per lattice site. All of these are semiconductors.
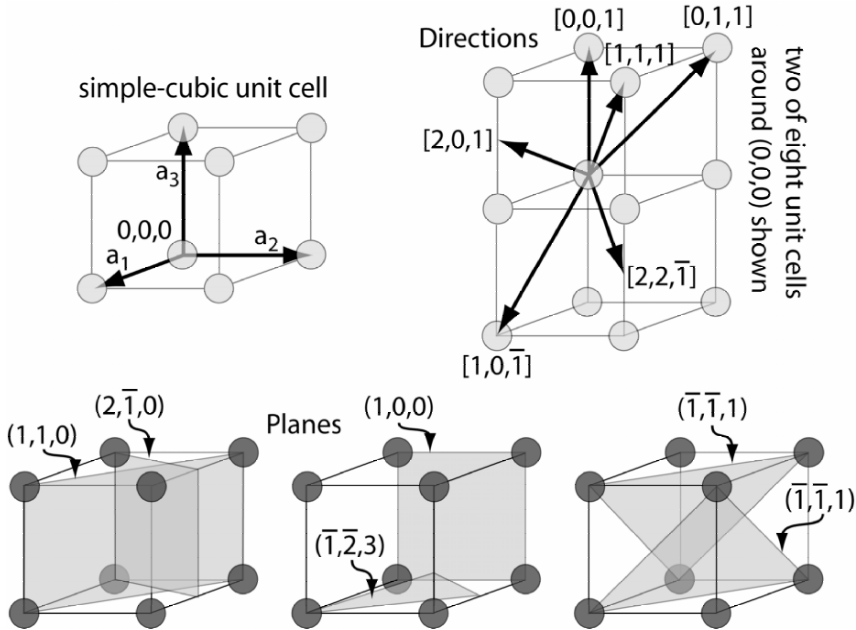
Figure 4.2: Shows the Miller indexing scheme for crystal planes and directions in a simple cubic crystal. Similar indexing methods are used in other Bravais lattices, although the basis vectors are not necessarily parallel to Cartesian coordinate axes. In hexagonal lattices an alternate labeling scheme employing four indices in which only three of the four are independent is often used. Notice that the two ($\bar{1}\,\bar{1}1$) planes marked are adjacent and parallel to one another.

Semiconductors can be found having many different crystal structures. Examples include silicon (a face centered cubic [fcc] lattice), cadmium sulfide (hexagonal close packed [hcp]), lead telluride (simple cubic), copper indium diselenide (tetragonal), and amorphous silicon (amorphous). Common semiconductors are built of sp^3 hybrid molecular orbitals with a tetrahedral geometry (for both hcp and fcc structures) as discussed in detail in Chapter 5. Lead Telluride and other IV-VI compound semiconductors are an exception to this behavior. This is not surprising when one considers that the IV-VI semiconductors have an average valence of five, while the common group IV, III-V and II-VII semiconductors have an average valence of four. Thus, the IV-VI semiconductors cannot produce an energy gap through tetrahedral bonding, as sp^3 hybrid orbitals can only accommodate four electrons per atom in bonding states.
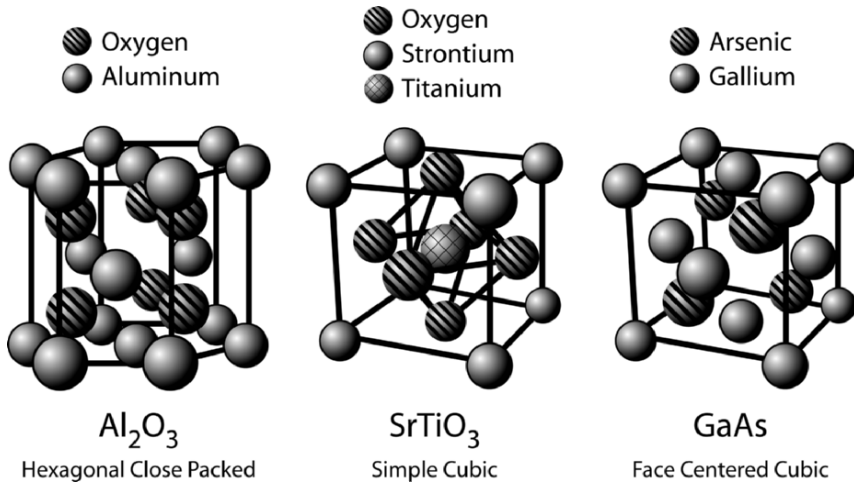
Figure 4.3: Shows the crystal structures of three compounds used in microelectronics. Each has a different Bravais lattice. To see the Bravais lattices of each compound, look only at the aluminum, strontium, and gallium atoms, respectively.

Chemical reactions giving rise to compounds place additional constraints on the arrangement of atoms. In general, there will be at least the number of atoms per Bravais-lattice structural-unit as there are atoms in the chemical formula or an integral multiple of the formula. Three compounds used in microelectronics illustrating three crystal structures are shown in Figure 4.3. The advanced dielectric $SrTiO_3$ has a simple cubic structure with Sr atoms at the cube corners, a Ti atom at the cube center, and O atoms at the face centers. In some ways this looks like an fcc structure with an extra atom at the cube body center or like a body centered cubic structure with extra atoms on the face centers. In the end, the only periodic structure is the simple cube because only the simple cube takes you from any atom to another of the same chemistry. Within each cube there are one Sr, one Ti, and three O atoms as expected from the formula. In the common semiconductor GaAs there are two atoms in the compound and two atoms per lattice site. This compound has a fcc Bravais lattice and each cube contains four formula units of the compound. A different unit cell must be defined to encompass only one formula unit. For each Ga atom, there is an As atom displaced by one quarter of the unit cell length along all of the three axes. In Al_2O_3 there are five atoms in the compound and five in the hexagonal unit cell. Because the aluminum planes defining the hexagonal volume of the solid contain a layer of aluminum in the vertical center of the unit cell, the structure is the close packed arrangement rather than the simple hexagonal structure.

Close packed structures are distinct from simple structures because rather than stacking atoms directly on top of each other along a given axis, the atoms stack in the
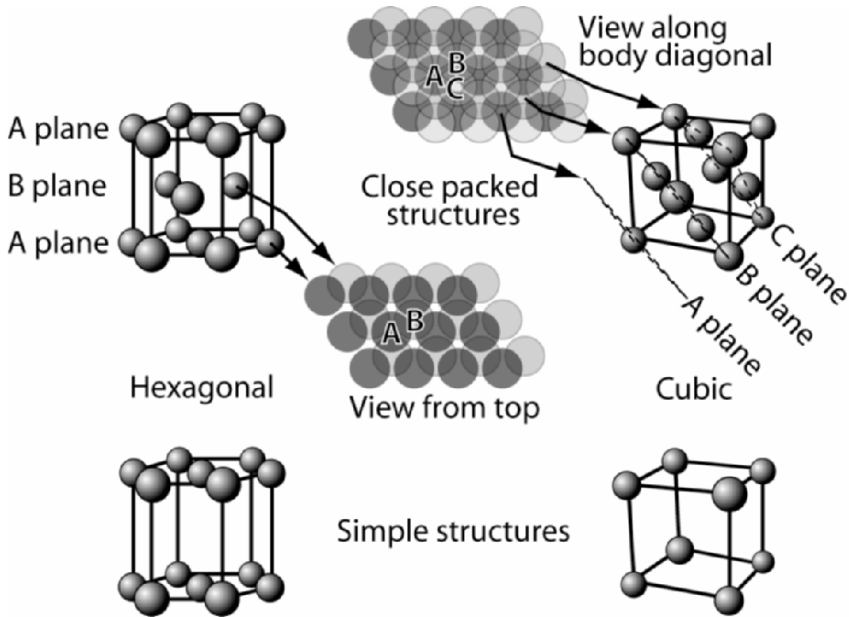
Figure 4.4: A comparison of fcc and hcp structures and their corresponding simple structures. Atoms in the "A" plane correspond to the darkest gray circles in the top view while atoms in the "B" and "C" planes are progressively lighter shades of gray.

hollows between atoms in a lower layer. The close packed face centered cubic (fcc) and and hexagonal close packed (hcp) structures are compared in Figure 4.4 with their non-close-packed cousins. If one considers the hexagonal planes that are the body diagonals of the fcc cube, one finds that there are three different planes displaced with respect to one another. In Figure 4.4 these are labeled A, B, and C. The hexagonal close packed structure consists of two such planes, A and B. The only difference between the fcc and hcp structures is that the fcc is formed by an ABCABC stacking of planes, while the hcp structure is stacked ABABAB with reference to the sites labeled A, B, and C in the figure. In other words, the fcc structure uses both distinct hollows between atoms in a given layer to stack the surrounding layers, while the hcp structure uses only one.

As noted above, a basic feature of all crystal lattices is that beginning from any point on the lattice and moving by any combination of translation vectors of that lattice you arrive at another lattice site. Diamond-structure semiconductors are composed of regularly stacked tetrahedra with atoms at their corners and at their centers. If one examines the possible lattices formed by arrangements of tetrahedra, one rapidly realizes that if one moves from a tetrahedron corner site to a tetrahedron center and then moves a similar distance along the same direction, one does not come back to

another atom. This means that bond directions from one atom to the next do not represent translation vectors of a lattice made up of tetrahedral subunits. Rather, the atoms in the corner and the center of the tetrahedra must lie on different sublattices within the Bravais lattice (see the GaAs structure in Figure 4.3). When a crystal structure contains two or more inequivalent atom sites (the two positions are not connected by lattice translation vectors) then the structure is said to have a multi-atom basis. No matter how one reduces the unit cell to its most fundamental repeat unit, one can never have fewer than two atoms in these structures. More complex compounds such as $SrTiO_3$ or Al_2O_3 have even more atoms in their structural unit cells. Note that atoms such as Si that crystallize in the diamond structure have two identical atoms on non-identical lattice sites. (The diamond structure is the same as for GaAs in Figure 4.3 if the Ga and As atoms are replaced by identical atoms such as Si.) The diamond structure is not itself a Bravais lattice. It is an fcc Bravais lattice with, for example, two atoms Si per lattice site. These two atoms have the same orientation with respect to one another at all sites of the fcc lattice. One can also view the structure as two fcc Bravais lattices displaced by 1/4 of a unit cell along each coordinate axis with respect to one another.

A number of semiconductors can exist in both the fcc and hcp lattices. A close examination of the two structures in Figure 4.4 shows that the primary difference is at the second-nearest neighbor distance. In the diamond (as in GaAs) or sapphire (Al_2O_3) structures shown in Figure 4.3 the distinction is even smaller as each lattice site has a pair of atoms associated with it so the distinction between the fcc and hcp lattices is essentially made at the fourth-nearest neighbor distance. Therefore, the energy difference between the two lattices and the difference in bonding will be very small. Of course, a small energy difference summed over a whole crystal can represent a significant energy. So long as the energy difference is greater than the thermal energy (k_BT per atom), the solid will still strongly favor the lower energy structure. However, factors such as temperature, impurities, or the presence of a lattice-distorting substrate can shift the balance and change the resulting crystal structure. Most of the II-VI semiconductors occur in both cubic and hexagonal forms although most have a preference for one or the other at room temperature.

Defects in these crystal structures are essential to determining the properties of the materials. The crystalline defects relevant to semiconductors will be discussed in detail in Chapter 7. Amorphous materials have no regular order so there are no well-defined defects in the material. Nonetheless, we will see in Chapter 8 that the continuum of distortions in the structures of amorphous semiconductors play a key role in determining their properties. Here we will list only the types of atomic-scale (point) defects in crystalline materials and leave more complex structures and detailed discussion to Chapters 7 and 8. Point defects in crystals include vacancies, interstitials, and antisites. Vacancies are missing atoms in the crystal structure. They are essential to the diffusion of atoms among lattice sites in many materials. Interstitials are atoms lying in spaces between atoms in the crystal structure. More open lattices such as the diamond structure accommodate interstitial atoms relatively

easily. Finally, in ordered structures one may disorder the atomic sites locally. For example, in GaAs there are a specific set of lattice sites for Ga atoms in an fcc arrangement and a separate set of fcc lattice sites for As. An antisite defect places an As on a Ga lattice site or vice versa. Extended line, planar and volume defects also occur but these disrupt the lattice and are left for description in Chapter 7.

4.1.2 The reciprocal lattice

It is important in understanding electron (and x-ray) diffraction to know how waves interact with solids. The energy band diagrams described in Chapter 2 all refer to electron wave vectors and diffraction of the electron wave off the crystal lattice at the Brillouin zone boundary. To understand diffraction one has to consider the lattice from the perspective of a wave, not from the perspective of points.

Any function, such as a function f(**R**) giving the probability of finding an atom at a given position **R** in real space, can be Fourier transformed to provide an indication of the amplitude f(**k**) of waves having wave vector **k** which, if summed or integrated, replicate f(**R**). Mathematically, the Fourier transform relationship is:

$$f(\mathbf{R}) = \int f(\mathbf{k})\, e^{i\mathbf{k}\bullet\mathbf{r}} dk . \qquad\qquad 4.1$$

The amplitude function f(**k**) is similar to a display of the intensity of sound as a function of wavelength on a graphic equalizer or to the intensities of colors in a spectrum produced by passing a light beam through a prism.

Perfect infinite crystals may be described by a series of harmonics of very well-defined wavelength, while a small crystal has much less capability to specify the waves needed to reproduce it exactly. Therefore, a much wider range of wavelengths is allowed. One can see this difference if we take two patterns of dots overlain upon one another. Twisting one set of dots relative to the other by a small angle is immediately obvious for a large dot array while it is not so detectable for a small array. This comparison is shown in Figure 4.5. The result is that a Fourier transform of a small perfect crystal gives significant intensity over a range of wave vectors, **k**, while a large crystal yields intensity only for much more specific values of **k**. The Fourier transform intensity spectrum f(**k**) directly determines the diffraction of waves off the lattice, as we will see shortly. Thus, a small crystal will give a broad diffraction peak in a diffraction spectrum, while a larger or more perfect crystal will give sharper diffraction peaks, as shown in the figure.

The relationship between the intensity spectrum, f(**k**), and diffraction may be seen by considering the conditions necessary for diffraction in more detail. In diffraction an electromagnetic or electron wave with flat wave fronts and wave vector $\mathbf{k}_i$ strikes an array of atoms as shown in Figure 4.6. The interaction of the wave with the electron clouds of the atoms causes them to emit spherical waves of the same wavelength.
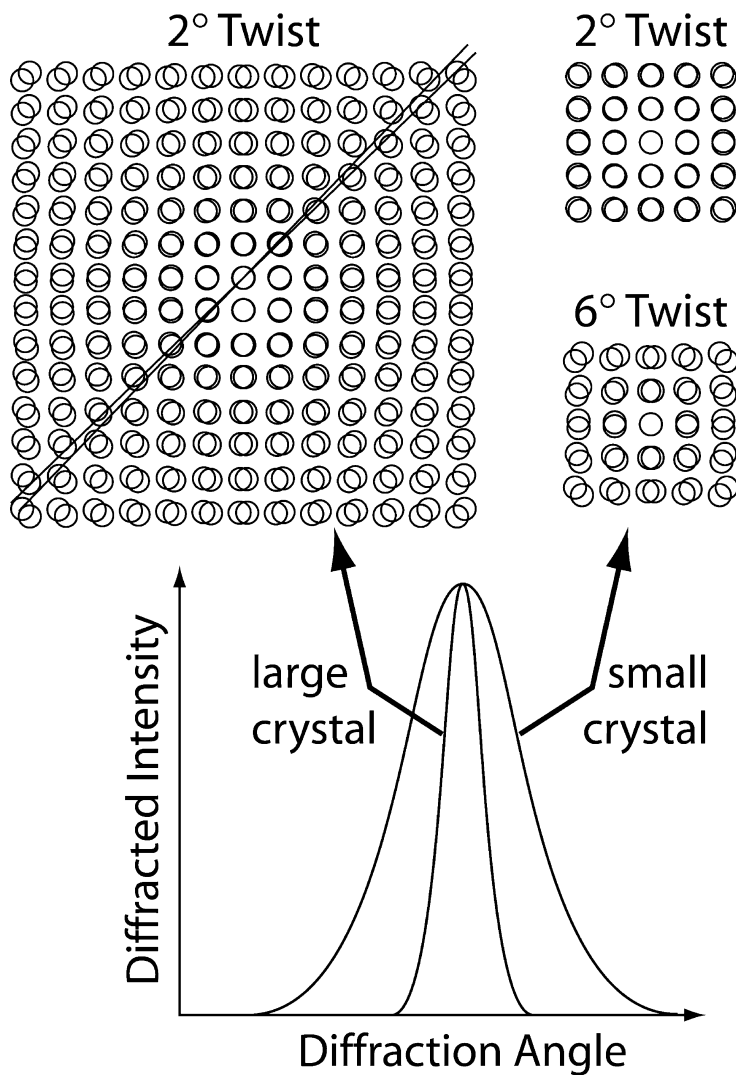
Figure 4.5: Shows the effect of crystal size on sensitivity to angular distortions. A "large" crystal shows an obvious effect of a 2° twist of one layer with respect to another. By contrast, a "small" crystal shows an almost undetectable effect. To achieve a similarly visible distortion in the small crystal a 6° twist is required. The sensitivity to variations in angle produces similar effects on diffraction patterns. Large crystals produce diffraction (constructive interference) only over a very small angular range. Small crystals allow diffracted intensity over a much larger range.
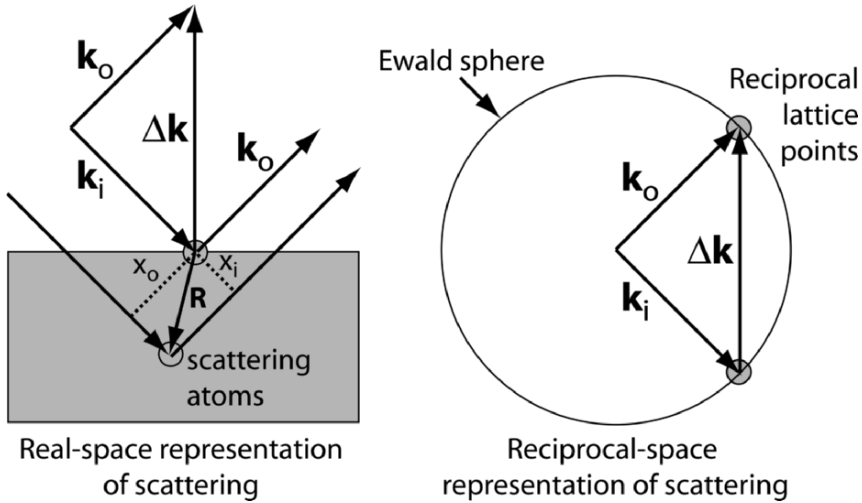
Figure 4.6: Shows the relationships among the incident, $\mathbf{k}_i$, and scattered, $\mathbf{k}_o$, wave vectors, the change in wave vector, $\Delta\mathbf{k}$, the extra path lengths, x_i and x_o, of one part of the scattered wave front along the incident and outgoing waves, and the distance between two scattering centers (atoms) in real space, $\mathbf{R}$. Also shown is the equivalent reciprocal space construction based on the Ewald sphere concept showing that diffraction occurs when $\Delta\mathbf{k}$ begins and ends on a reciprocal lattice point.

As these waves interfere, they produce an outgoing wave at a particular angle having a wave vector $\mathbf{k}_o$. If no energy is absorbed in the process (which is typically the case in diffraction), the magnitude of the two wave vectors are the same, but their directions are different. To determine at what angle interference will be constructive one may ask what the extra path length is for the incident wave to strike an atom at position $\mathbf{R} = u\mathbf{a_1} + v\mathbf{a_2} + w\mathbf{a_3}$, relative to an atom at an arbitrary reference position, $0,0,0$ where $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ are the primitive translation vectors for the lattice. The extra distance for the incident wave is $x_i=\mathbf{k}_i\cdot\mathbf{R}$ and the extra distance for the outgoing wave is $x_o=\mathbf{k}_o\cdot(-\mathbf{R})$. The total extra distance is then $x=x_i+x_o=(\mathbf{k}_i-\mathbf{k}_o)\cdot\mathbf{R}$. It is convenient to define $\Delta\mathbf{k}=\mathbf{k}_i-\mathbf{k}_o$. Then, constructive interference occurs when $x=\Delta\mathbf{k}\cdot\mathbf{R}$ is an integer multiple of 2π. If we write $\Delta\mathbf{k}=h\mathbf{b_1}+k\mathbf{b_2}+l\mathbf{b_3}$ and if we define a relationship between the $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ and $\mathbf{b}_1$, $\mathbf{b}_2$, and $\mathbf{b}_3$ vectors as

$$\mathbf{b}_1 = (\mathbf{a}_2 \times \mathbf{a}_3)/(\mathbf{a}_1 \bullet \mathbf{a}_2 \times \mathbf{a}_3)$$
$$\mathbf{b}_2 = (\mathbf{a}_3 \times \mathbf{a}_1)/(\mathbf{a}_1 \bullet \mathbf{a}_2 \times \mathbf{a}_3),$$
$$\mathbf{b}_3 = (\mathbf{a}_1 \times \mathbf{a}_2)/(\mathbf{a}_1 \bullet \mathbf{a}_2 \times \mathbf{a}_3)$$

4.2

then the dot product becomes $\Delta\mathbf{k}\cdot\mathbf{R}=2\pi(uh+vk+wl)$. If u, v, and w are integers based on the definition given above for the crystal lattice atom positions, then this imposes the additional criterion that h, k, and l also must be integers. This implies that $\Delta\mathbf{k}$ must fall upon a lattice defined by points $h\mathbf{b}_1$, $k\mathbf{b}_2$, and $l\mathbf{b}_3$.

> **The reciprocal lattice is defined by the set of points** $h\mathbf{b}_1$, $k\mathbf{b}_2$, and $l\mathbf{b}_3$ **with the basis vectors defined by Equation 4.2.**

The diffraction condition may also be described by the "Ewald sphere" construction, shown in Figure 4.6. Vectors $\mathbf{k}_i$ and $\mathbf{k}_o$ have the same length and therefore can be presumed to define the surface of a sphere if both originate from the same point. Positioning the end of vector $\mathbf{k}_i$ on a reciprocal lattice point, diffraction will result in any direction in which the sphere passes through another reciprocal lattice point. The angle between $\mathbf{k}_i$ and $\mathbf{k}_o$ intersecting these two reciprocal lattice points, (0,0,0) and (h,k,l) defines the diffraction angle for reflection off the real-space lattice planes (h,k,l).

The conditions described here also define the conditions for diffraction of electron waves at the Brillouin zone boundaries. Likewise the Brillouin zones described in Chapter 2 are reciprocal-space objects with the symmetry of the reciprocal lattice rather than the real-space lattice. The reciprocal lattice points in Figure 2.5, for example, are located at points $h\mathbf{b}_1$, $k\mathbf{b}_2$, and $l\mathbf{b}_3$. The reciprocal lattice for a simple cubic system with basis vectors $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ has reciprocal lattice vectors parallel to the real space vectors. However, larger distances in real space correspond to shorter distances in reciprocal space. Thus, planes that are widely spaced in real space have closely spaced reciprocal lattice points and vice versa. One may determine by examination of Figure 4.2 that the (100) planes are $\sqrt{3}$ times farther apart than are the (111) planes. In general, the distance, d, between (hkl) lattice planes in a cubic system may be shown to be:

$$\frac{1}{d^2} = \frac{h^2 + k^2 + l^2}{a^2}, \qquad\qquad 4.3$$

where a is the cube side length.

4.2 BASIC THERMODYNAMICS OF MATERIALS

An essential issue in materials science is to establish the stability of any of the crystal structures discussed above relative to any other, to determine the energy released or absorbed upon reaction to form or decompose compounds, and to determine how conditions such as temperature and pressure affect the stability of materials. This section describes the basic approach to tackling these questions.

Before calculating the stability of a material, we need to decide how to describe it in a well-defined way. We will make extensive use of the term "phase."

> **A phase is a chemically and structurally homogeneous material.**

Thus, liquids and solids are different phases. Solids of different crystal structures are different phases even if they are chemically the same. Mixtures of two solids are called two-phase mixtures. Alloys are single-phase mixtures because one cannot distinguish isolated regions of one or the other constituent within the alloy. Thus, a phase can have different chemical compositions in different mixtures and still be the same phase as long as a given mixture is homogeneous and no structure change occurs. For example, varying amounts of sugar can be dissolved in water. As the mixture contains more sugar it becomes thicker and more syrupy. However, it is still the same phase (a chemically and structurally homogeneous liquid sugar water solution). What crystal structure forms for a given solid is a question of what phase is most stable.

When a material changes its crystal structure or when a solution of several constituents decomposes into separate parts with different compositions or structures, this is known as a phase change. Phase changes occur with changes in temperature, pressure, or composition as a result of a change of most stable structure. However, a phase change does not generally occur all at once across an entire material. (The martensitic transformation in steel is about the closest thing to an exception to this statement where the phase transformation moves across the material at the speed of sound and requires no long-range transport of atoms.) A phase change requires formation of a distinct interface between initial and final phases. Although the phase change might be favorable if it could be accomplished across the entire solid, one of the primary questions in determining whether a phase change will occur is how hard it is to form such an interface. Furthermore, it is often necessary to redistribute atoms if the product phases have different compositions. This atom movement takes time. Two fundamental fields describing these topics are thermodynamics and kinetics. Thermodynamics answers the question: which state does the system prefer, while kinetics asks: how fast will the system transform from an initial unstable state to any other state.

The stability of any material or interface is determined by the energy (or chemical potential) of the starting state relative to the energies of potential reaction products or other crystal structures. From our perspective, the key components in determining the energy of a material are the heat of the material, the work done in forming it, and any chemical bond energies. These contributions determine the total energy of the solid as a function of system variables such as temperature, pressure, and number of atoms in the system.

Let us examine these components from a thermodynamic perspective. In Chapter 2 (and again in Chapter 5) we consider quantum mechanical methods for determining

the energy change upon bond formation and hence the energy of a solid at zero temperature. Knowing the vibrational modes allows us to go on to estimate the energy at a non-zero temperature. This is fine if we have the patience and determination to work out the full quantum mechanical results for all arbitrary configurations of the system to calculate which is most stable. A more traditional and practical method is to look at the response of a material to macroscopic variables such as temperature and pressure and their effects on the energy of the material. From these results we can work out a lot about the energies of the phases.

Thermodynamics begins by relating the energy of a material to the work, W; heat, Q; and bond energy Q_B that are created or absorbed in its formation. Other energies could be included as necessary. For example, the gravitational energy is not normally important to a chemical reaction and is thus generally ignored. Thus, the total energy, E, of a solid may be written

$$E = Q + W + Q_B + \text{other terms as needed.} \qquad 4.4$$

Likewise, when a system changes state, its total energy is modified based on heat and work done during the transformation and changes in bond energy as,

$$dE = dQ + dW + dQ_B + \dots \qquad 4.5$$

We will consider the heat, work, and bond energies of forming a solid separately.

The work done in creating or modifying a material can be divided into a number of contributions including mechanical work, electrical work, surface work, etc… The mechanical work done, dW, in any process is

$$dW = -PdV, \qquad 4.6$$

where P is the pressure on the material and dV is the change in volume resulting from the process by which the change occurs.

If one considers a collection of atoms widely separated such that they have virtually no interaction energy and one allows the atoms to come together to form a solid, the system converts potential energy to kinetic energy. In other words, to remain at constant temperature it must do work on some outside environment as it comes together to release energy. Typically, the amount of chemical energy produced is the cohesive energy of the solid and is the chemical work of bond formation, dQ_B. The magnitude of the cohesive energy is discussed in more detail in Chapter 5. It can also be expressed as the chemical potential of an atom, μ, multiplied by the number of atoms, N, in the system. Because there is a separate chemical potential for each chemical species in the system, one must sum over all i species present:

$$dQ_b = \sum_i \mu_i dN_i. \qquad 4.7$$

The heat of formation of a material dQ results from a change in the number of states accessible to it at a given temperature. The number of states at a given energy is

related to a quantity called the entropy of the system, S. By definition, $S = k_B \ln\Omega$, where Ω is the number of states accessible to a system with an energy between E and E+dE.

To get a better sense of entropy, consider a single gas molecule in a large volume. That molecule will have a large number of energetically-equivalent places it can be within the volume and consequently a high entropy. Simply confining the same molecule in a smaller volume without otherwise changing the system reduces the number of places it can be and hence reduces the entropy of the molecule. A characteristic of thermodynamics is that spontaneous processes increase the total entropy of the system. All reversible processes involve no change in the total entropy of a system.

The nature of temperature is also connected to this concept of the number of states accessible to a system. Specifically, the temperature of a system is $T = \partial S/\partial E$ by definition. In other words, the temperature of a system is the change in the number of states accessible to a system resulting from a small change in the energy of the system with all other system variables held constant. One can rewrite this relationship as dE = TdS when other variables are held constant. A quick inspection of Equation 4.5 reveals that holding volume and number of atoms constant $(dV=dQ_B=0)$, implies that

$$dQ = TdS. \qquad\qquad 4.8$$

from which we can see that the generalized change in energy of a system for a constant number of particles is:

$$dE = TdS - PdV. \qquad\qquad 4.9$$

If we revisit the issue of gas molecules in a volume, we can see that the change in entropy of an isolated system at constant temperature and number of molecules would be directly related to the change in volume as dS = P dV/T.

Equation 4.9 is fine as long as the number of particles in a system is kept constant. There is an obvious problem if this is not the case. For example, suppose that we keep the volume and energy of a system constant but allow the number of particles to change. If the temperature and pressure are not allowed to change, then the entropy must change as the number of states accessible to the system is modified. Alternately, changing the number of particles at constant volume and entropy must change the energy of the system as each atom added also adds a unit of chemical potential of that species. To account for changes in the number of particles of type i:

$$dE = TdS - PdV + \sum_i \mu_i dN_i. \qquad\qquad 4.10$$

The index i can also designate a series of phases.

Returning to the absolute energies of Equation 4.4, rather than energy changes, one can write

$$E = TS - PV + \sum_i \mu_i N_i.$$ 4.11

Several other energies are useful to define. The enthalpy, H=E+PV describes the energy of the solid excluding mechanical work and describes the energies associated with bond energy and configurational energy. The Helmholtz free energy, F=E-TS removes contributions due to entropy, and the Gibbs free energy is G=F+PV. From these definitions,

$$H = TS + \sum_i \mu_i N_i$$

$$F = -PV + \sum_i \mu_i N_i.$$ 4.12

$$G = \sum_i \mu_i N_i$$

Each of these free energies is most useful for describing certain behaviors of the system or for conditions under which only certain variables are changing. For example, because it is directly related only to the chemical potential of a set of particles, the Gibbs free energy is a broadly useful quantity for determining the chemical stability of a closed system of given size. Furthermore, because systems are in equilibrium when their chemical potentials are equal, the condition of equilibrium is when there is no change in their Gibbs free energies if they exchange particles. An example of a system that can exchange particles is a mixture of two phases in a solid. Phase changes will occur whenever the chemical potential of one phase is lower than the other. Stability is therefore the condition in which the chemical potentials of all phases or particles are equal, as was discussed in Chapter 2.

One final point to note about the energy of a system is that at higher temperatures the entropy term becomes more important than the bond energy terms. Therefore, as one heats a solid one often observes changes in structure because different crystal lattices have different entropies. If one makes mixtures of materials the entropy of mixing often becomes important. It is the desire to mix materials to raise the entropy of the system that produces a corresponding decrease in the melting points of many mixtures. This is how antifreeze works.

4.3 PHASE DIAGRAMS

It will be important in discussions in other parts of this text to be able to describe easily which phases will be present under a variety of conditions and if a combination of phases is in equilibrium or not. For example, at what temperature will an element, alloy or mixture of phases melt? What is the composition of the material that melts at
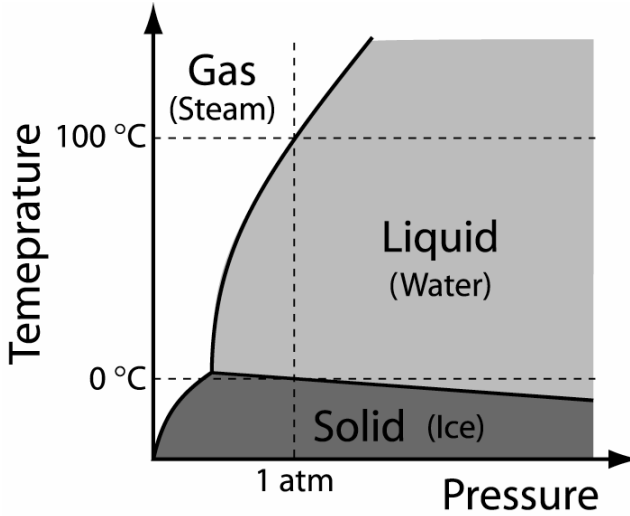
Figure 4.7: The pressure-temperature phase diagram for water. Note the "triple point" which occurs just above 0°C at a fraction of an atmosphere pressure where solid, liquid and vapor can coexist. This must be a point because all degrees of freedom are taken up by the three phases present simultaneously. This can only occur at a single point.

such a temperature? And so forth. These questions can be answered immediately by reference to a phase diagram. This is a map showing which phases are stable as a function of variables such as temperature, pressure, or composition.

It can be shown that the number of degrees of freedom *F* which a system has available to it is

$$F = C - P + 2, \qquad\qquad 4.13$$

where *C* is the number of constituents in the system, *P* is the number of phases, and the 2 refers to the availability of work and heat as additional variables. Normally the laboratory variables we use to affect work and heat are temperature and pressure, although entropy and volume could also be used. If we have a single constituent, such as water, the number of degrees of freedom will be *F*=3-*P*. Therefore, we can vary temperature and pressure independently in the presence of a single phase or we can change either temperature or pressure if we want to have two phases present simultaneously (one degree of freedom). This means that the boundary between two phases (where they coexist) will be a line on a temperature-pressure diagram. A typical example of a phase diagram is that for water, shown in Figure 4.7. It is even possible to have three phases coexist as long as it is only at one temperature and pressure. This is known as the "triple point."
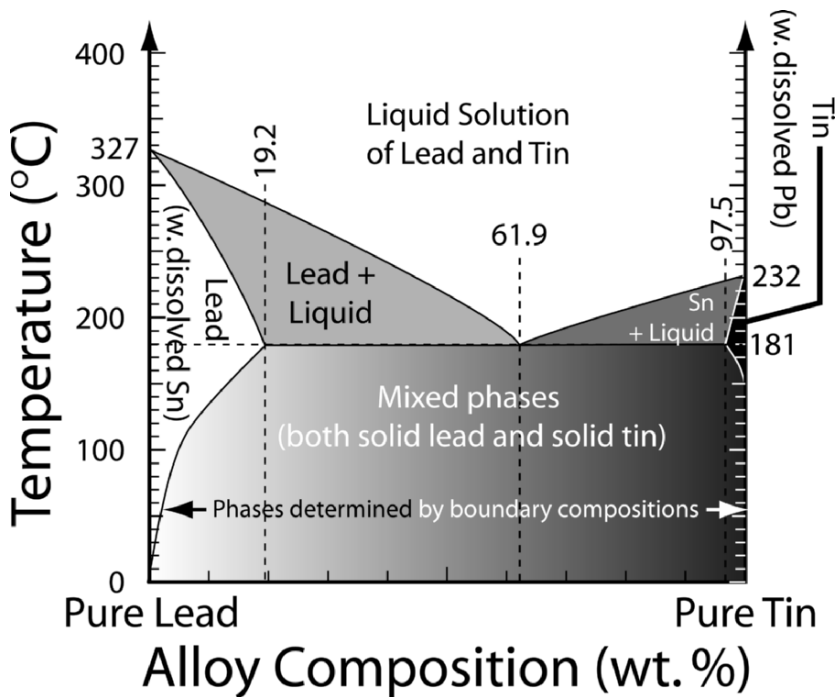
Figure 4.8: The Pb-Sn binary phase diagram. This type of phase diagram is typical of materials with limited solid solubility.

When two components are present, an additional degree of freedom exists. The three degrees of freedom are now temperature, pressure and composition. This means that two phases can be present in a two-dimensional region of phase space. One can, for example, change both the composition of the mixture and the temperature and retain the two phases. It is even possible to have three phases present together. For example, two solid phases and a liquid phase can coexist, but only on a line.

A typical example of a phase diagram with coexisting phases is a binary eutectic system such as that of Pb and Sn, commonly used in solders. The Pb-Sn phase diagram is shown in Figure 4.8. There are a number of features to notice. First, both Pb and Sn can dissolve some of the other element in their solid form, but Pb can dissolve much more Sn and this solubility persists to much lower temperatures than for solubility of Pb in Sn. Second, at temperatures below 181°C (the horizontal line on the figure) the system consists of a mixture of solid phases of Pb and Sn or a solid

solution of one or the other of these. The composition of the phases is determined by the left and right boundaries of the two-phase region at the temperature in question.

There is a "triple point" at which a liquid phase (in the case of Pb-Sn, a liquid solution of Pb and Sn), and two solid phases coexist. This is known as the eutectic point. It occurs at a specific composition, the eutectic composition, and melts/freezes completely at a specific temperature, the eutectic temperature. In the case of Pb-Sn, the eutectic composition is ~61.9 weight % Sn, and the eutectic temperature is 181 °C. Notice that the eutectic solid is a two-phase mixture. In spite of this, both phases melt at the same temperature for the eutectic alloy. The maxima of solid solubility for Sn in solid Pb and Pb in solid Sn occur at the eutectic temperature and are, 19.2 weight % and 97.5 weight % Sn, respectively. A solid mixture having the eutectic composition on average is the only mixture that melts entirely at a single temperature. All other compositions (except pure elements) in a eutectic system partially melt leaving a solid with altered composition (if kinetics allow the solids to maintain equilibrium) in contact with a liquid of different composition.

For example, a Pb-Sn alloy containing 19.2 weight % Sn begins to melt at 181°C, producing a liquid with the eutectic composition. Because the liquid has much more Sn than the starting solid composition, the solid is slightly depleted of Sn as the liquid forms. As the temperature rises further, more liquid forms and the solid is more depleted of Sn. However, because of the rate of creation of liquid and loss of solid, the liquid must also become progressively depleted in Sn. The two phases present have compositions given by the boundaries of the two-phase region on the phase diagram at a given temperature. Thus, to decide on the compositions of the liquid and solid in equilibrium after melting the 19.2 weight % Sn alloy one draws a horizontal line across the phase diagram at the temperature in question. Where this line intersects the boundaries of a two phase region determines the phase compositions. For example, at ~263°C the two phases in equilibrium for starting Sn fractions below the eutectic composition will be solid Pb with about 10 at.% Sn dissolved in it and a liquid of about 30 at.% Sn. The entire mixture has melted at ~289°C.

Often mixtures of two or more elements form compounds – in some cases a wide variety of compounds. Taking the Si-Ta phase diagram shown in Figure 4.9 as an example one can observe a wide range of behaviors. First there are seven separate phases represented on this diagram including the liquid phase and the two elemental phases, Si and Ta. There are four intermetallic compounds. Because these compounds require significant space to write as chemical formulas, they are usually designated with small Greek letters. Most of the phase diagram consists of two-phase regions and none of the phases except the liquid can deviate much from a single composition.
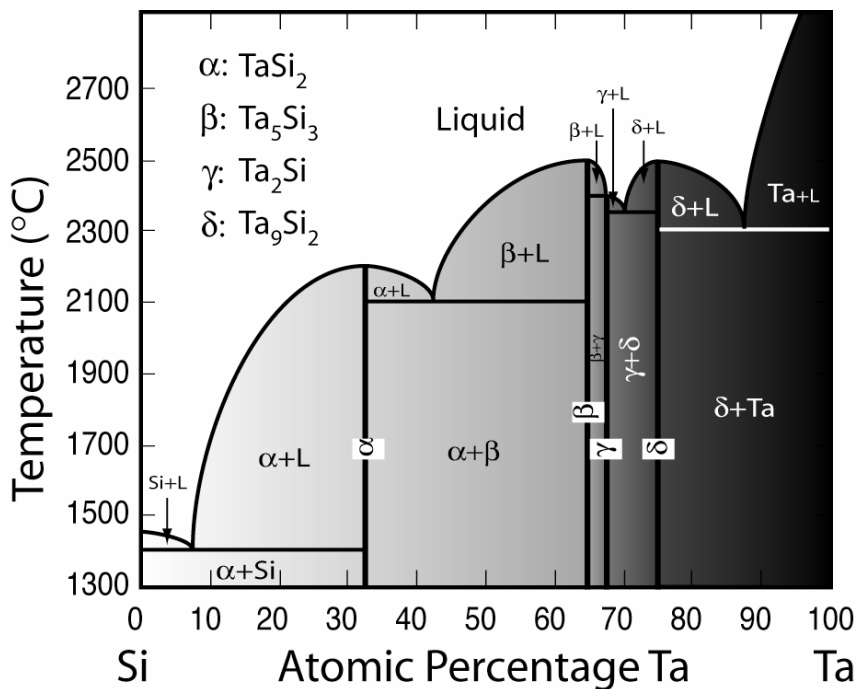
Figure 4.9: The Si-Ta phase diagram. Each vertical line corresponds to a separate compound and is labeled with a Greek letter. Note that between each pair of compounds a simple binary phase diagram can be drawn. Phase diagrams between specific compounds may be referred to as pseudobinary phase diagrams.

Limited solid solubility in the elemental phases is not uncommon when compounds form easily, as the presence of these compounds indicates a driving force in favor of clustering and ordering of the atoms. Between any two compounds one finds a simple binary phase diagram behavior. One may therefore draw a "pseudobinary" phase diagram representing the equilibrium between the compounds bounding each region. Most of the individual equilibria between solids in this phase diagram are eutectic-like. Note that a liquid having exactly the composition of one of the compounds freezes to the single phase but all other liquids form two solid phases when frozen. Also notice that although the solids have very little solubility in each other, the liquid is one phase across the whole composition range. Thus, one would notice a continuous change from liquid Ta to liquid Si in the single phase melt as the composition changed.

The above discussions refer to phase equilibria for unary (water) and binary (Pb-Sn, or Ta-Si) materials. Very few semiconductor alloys exist as simple binaries. An example of one that does is the Si-Ge alloy, discussed in detail in Section 6.4. Most semiconductor alloys contain at least three elements. Therefore, for a phase diagram there should be three possible composition ratios to account for. However, because the sum of all atomic percentages must be 100%, only two free variables exist. Consider, for example, the AlAs-GaAs alloy. Increasing the Ga content at constant As level must decrease the Al in the alloy. The resulting composition space can be plotted effectively as points within an equilateral triangle. Each edge of the triangle represents an individual binary alloy, similar to those shown in Figures 4.8 and 4.9. A ternary phase diagram for the Al-Ga-As system is shown in Figure 4.10.

In Figure 4.7 we had a phase diagram for a single compound. In the plane of the plot we could therefore accommodate two variables, temperature and pressure, which determined the phases present. Adding a composition variable in a two-dimensional plot, such as those in Figures 4.8 and 4.9, we had to eliminate one variable. The most common case is to assume atmospheric pressure. Therefore, the binary phase diagrams have axes of composition and temperature at a fixed pressure. In Figure 4.10 and other ternary diagrams we now have two different compositions to represent – group III element composition, and group III to group V elemental ratio in the example case. Therefore, we have eliminated the last environmental variable. Figure 4.10 had to have been made at a fixed temperature and pressure (normally atmospheric pressure). Fortunately, most ternary systems of interest have a critical or particularly interesting temperature at which, for example, growth of crystals occurs or heat treatment of the solid takes place. Furthermore, the phase diagrams for semiconductors often remain relatively constant over a fairly wide temperature range.

Let us consider the ternary phase diagram in Figure 4.10 in more detail. To begin to evaluate it we note that the interior of the triangle contains no compounds. In other words, no compound such as $AlGaAs_2$ exists which would have a plotted position away from the edges of the triangle. Along each edge one may also search for compounds. Only two exist in this phase diagram, AlAs and GaAs. As both of these compounds have virtually no solubility for point defects, they occur at single compositions and are plotted as points on the phase diagram. There are no Al-Ga intermetallic compounds. Below we will consider a more complex case with phases within the triangle. All together there are five separate phases on this diagram, Al, Ga, As, AlAs, and GaAs. Note that this diagram shows no specific temperature and ignores both liquid metals and gaseous elemental As. These are all likely to exist as Ga, for example, melts just above room temperature. Furthermore, to be in equilibrium at any given temperature above 50% As, the system must be enclosed and under sufficient As pressure to keep the excess As from boiling. Consideration of vapor pressure is essential in any system including a high vapor-pressure compound or element. Therefore, we can expect that the diagram in Figure 4.10 must be somewhat schematic or must imply a closed system and ignore molten phases or
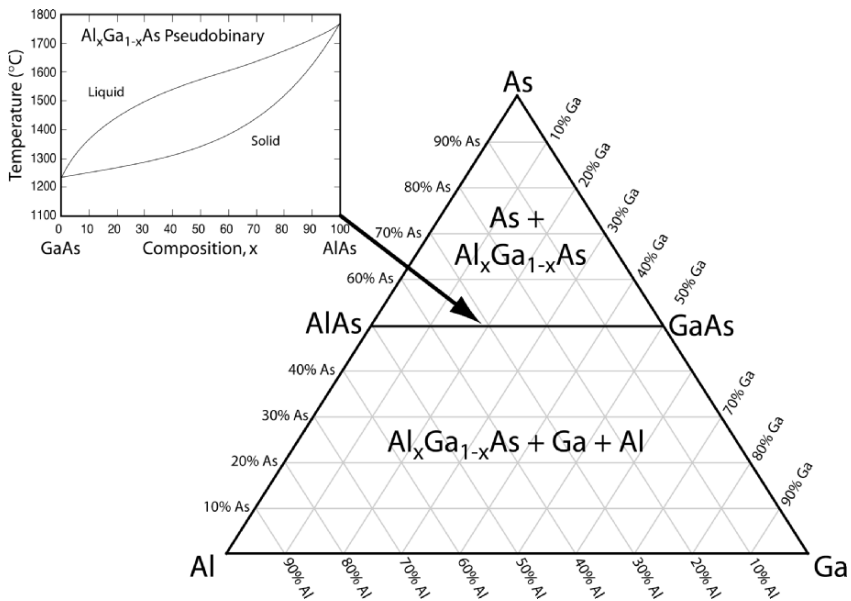
Figure 4.10: The ternary phase diagram for Al-Ga-As. Each light line represents alloys of constant composition as labeled for each line. The dark horizontal line represents the pseudobinary alloy $Al_xGa_{1-x}As$. This is a continuously-soluble alloy as shown in the inset. Figure inset based on data from Reference 1.

must represent a low temperature. It is common to simplify ternary phase diagrams in this way to present the general aspect of the phase relationships.

One can create a mixture of any pair of compounds and/or elements. In general there are three possible results: 1) you get a two phase mixture consisting of distinct regions of the compounds you mixed, 2) the two materials can dissolve in one another forming a single phase alloy, or 3) a reaction may occur, resulting in new phases. All of these situations may be found for Figure 4.10. Mixing AlAs and GaAs results in an $Al_xGa_{1-x}As$ alloy with no other phases (case 2). If one mixes As with GaAs, AlAs, or any alloy thereof, one gets a two-phase structure (case 1). Finally, mixing Ga with AlAs or Al with GaAs can result in reactions producing an $Al_xGa_{1-x}As$ alloy with both Ga and Al metallic second phases. In other words, one begins with two phases and ends with three (case 3).

Ternary phase diagrams have lines drawn connecting any two phases, which, when mixed, result in either case 1 or 2, above. These are referred to as "tie lines". Rephrasing, lines on the phase diagram indicate phases which, when mixed, will be unaffected or will alloy. The two-phase result indicates that the two materials are
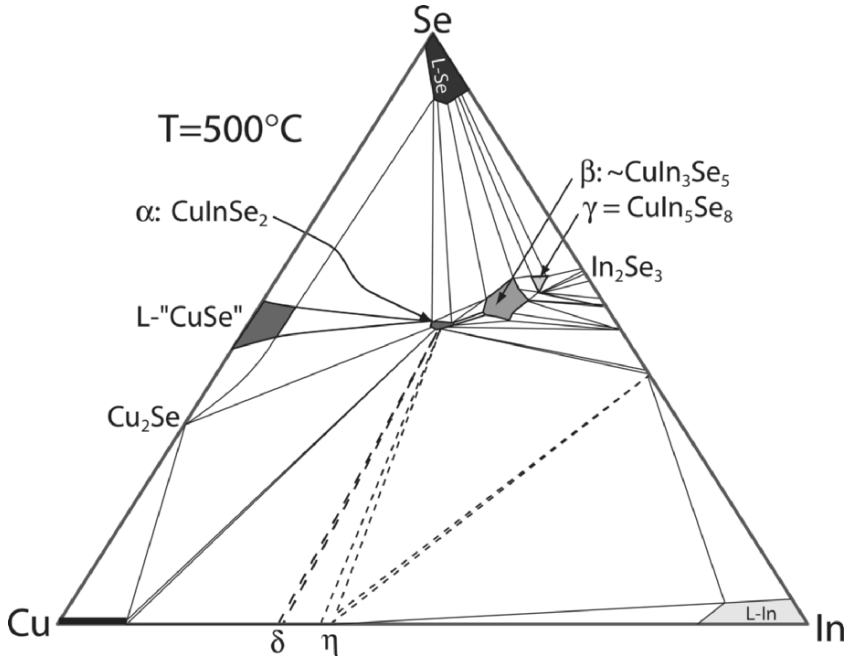
Figure 4.11: A ternary phase diagram for Cu, In, and Se at 500°C based on data in [2]. Additional data available for 750°C in [1]. Uncolored regions on the diagram are two or three phase areas. Numerous In-Se compounds occur which are not labeled to simplify the diagram. The diagram is approximate and most relevant for 500-750°C.

thermodynamically stable with respect to one another. This is desired for producing sharp interfaces such as at semiconductor-metal contacts. Any tie line may be represented as a pseudobinary alloy between two compounds or elements and a phase diagram such as in Figure 4.8 can be created. An example for the $Al_xGa_{1-x}As$ alloy is shown as an inset in Figure 4.10.

To illustrate some of the complexities possible in ternary phase diagrams, consider the Cu-In-Se ternary shown in Figure 4.11. This ternary system contains three semi-conducting phases within the triangle at 500°C, $CuInSe_2$, $CuIn_3Se_5$, and $CuIn_5Se_8$. Note that each of these is represented by a shaded area rather than a point because each can dissolve a significant number of defects and produce the same phase over a range of compositions. Numerous additional phases are present on the perimeter of the diagram including some with extended solid solubility. For example, a very In-rich liquid solution containing some Cu and Se occurs near the In-corner of the diagram. The phases present in a given region of the diagram will be the phases

defining the corners of that region. Liquids generally form solutions and are single phase, as occurs in this diagram.

All three phases within the triangle lie roughly along a line connecting Cu_2Se and In_2Se_3. Thus, it is common with this system to discuss a tie line between these two compounds. However, a more precise description would be to refer to a series of tie lines. In some instances such as between $CuIn_3Se_5$ and $CuIn_5Se_8$, rather than a tie line, a volume of the phase space connects the compounds in which a range of compositions occur. Similar behaviors are found in many semiconductor quaternary alloys as will be seen in Chapter 6. The various lines and dashed lines (for estimated boundaries) indicate tie lines connecting phases. Many of these lines show stable mixtures of materials. For example, vacuum growth of Cu-rich $CuInSe_2$ results in the formation of Cu_2Se second phases, usually on the surface of the $CuInSe_2$, in agreement with the diagram. The complexity of ternary phase diagrams can make them difficult to interpret and difficult to plot. However, they are very convenient for determining the thermodynamic stability of interfaces.

Fortunately, when dealing with semiconductors, we are more often faced with phase diagrams such as the Al-Ga-As ternary in Figure 4.10, which can be reduced to a few relevant binaries, rather than complex ternaries such as in Figure 4.11. By contrast, metallic contacts to semiconductors usually require looking at the more complex ternary phase diagrams to determine reaction and stability issues. A very complex situation such as the Au-Ge contact to $Al_xGa_{1-x}As$ would, in principle require a five element phase diagram. This is impossible to represent even with a three dimensional plot, so one must be content with an estimation based on, for example Au-Ga-As, Au-Al-As, Ge-Ga-As and Ge-Al-As ternary diagrams supplemented as possible by experiments.

4.4 KINETICS

Chemical reactions and phase transformations occur between elements or materials because by organizing or reorganizing atoms, the energy of the electrons in the system (or the chemical potential) can be lowered, in the case of an exothermic reaction, or raised if that energy can be obtained from a temperature bath (endothermic reactions). Whether a reaction occurs and what the energy barriers are is a question of thermodynamics. However, even exothermic reactions may never occur because the atoms need to move relatively long distances under conditions where movement is prohibitively slow. The question of how fast a reaction will occur is one of kinetics rather than thermodynamics. A common example of a kinetically-limited reaction is the stability of metals in air. All metals except Au will oxidize exothermically. Al reacts particularly rapidly with oxygen. Indeed, it is the fuel in many solid rocket boosters. What prevents Al soda cans from bursting into flame is kinetics.

Getting a reaction to proceed usually requires the following 1) sufficient reactants to provide the source materials for the reaction, 2) a small enough number of product molecules being present to prevent reverse reactions from dominating in reversible reactions (reversible reactions have little or no entropy change during the reaction), 3) sufficient atomic mobility that the atoms can sample the configurations of the system in which products are formed, 4) the energies of intermediate states between the reactants and the products are not too high, and 5) no parasitic (unintended) reactions occur that will dominate in competition with the desired reaction. All of these can be used to advantage or can represent a problem in a given process.

4.4.1 Reaction kinetics

For reversible reactions, reactant and product concentration effects are often modeled with the Law of Mass Action, which states that the reaction rate is proportional to a ratio of the concentrations of reactants to products, each raised to an appropriate power. Consider the reaction: $2A + BC_2 \rightarrow A_2B + C_2$ where A and BC_2 are the reactants and A_2B and C_2 are the products. Suppose that the concentrations of these species are [A], [BC_2], [A_2B], and [C_2], respectively. The rate of the forward reaction is:

$$R_f = k_f \frac{[A]^2[BC_2]}{[A_2B][C_2]},$$

4.14

where the forward reaction rate constant is

$$k_f = k_0 \, e^{-\Delta H / k_B T},$$

4.15

and k_0, and ΔH are the attempt frequency and the energy barrier for the reaction. The energy of the system during a typical exothermic reaction is shown schematically in Figure 4.12. Because the reaction requires two A atoms, it is necessary to bring them together simultaneously. (Otherwise the reaction would have to be broken down into two sequential reactions with one A atom each.) The time required to find two A atoms at adjacent locations at the same time is proportional to $[A]^2$. If we had required three A atoms the concentration would have been cubed. The reverse reaction rate is the inverse of the forward rate, leading to a change of sign of ΔH and the switch of reactants and products. In general the value of k_0 will also change. At high temperatures where the reaction rate constants, k, are all large, the forward and reverse reaction rates are determined primarily by the ratio of reactants to products. If [A] and [BC_2] are much larger than [A_2B] and [C_2], then the forward reaction rate will be high and the reverse rate will be low. When the converse is true, small [A] and [BC_2] relative to [A_2B] and [C_2] gives a predominant reverse reaction. If the temperature is low then the exothermic reaction rate will tend to dominate over the endothermic rate because of the difference in the values of the forward and reverse attempt rates, k_f and k_r.
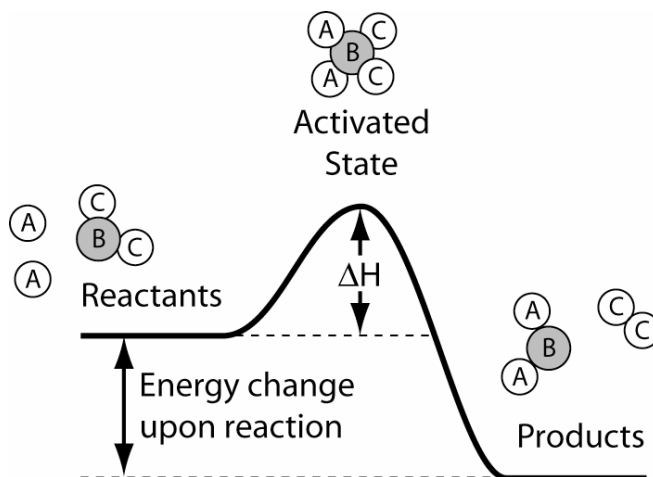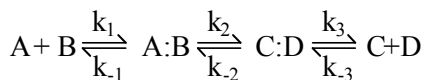
Figure 4.12: The energy of a system of atoms during the course of a reaction such as 2A + BC$_2$ → A$_2$B + C$_2$. The reactants must be configured for reaction. In the example this might mean attaching two A atoms to the BC$_2$ molecule prior to reaction. The reverse reaction would cross a higher energy barrier climbing the entire energy hill from the right ("Products") to the activated state energy.

In the event that there are multiple steps in the reaction, the rate constants of the various steps can be included separately in an overall rate. For example, consider a three-step reaction in which two reactants form a complex, the reaction occurs and then the complex decomposes to yield products:

$$A + B \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} A{:}B \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} C{:}D \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} C + D$$,

where A and B are reactants, C and D are products, the formation step for the A:B complex has a rate constant k$_1$, calculated based on Equation 4.15, and so on. The forward reaction rate is k$_2$, and the product complex decomposition rate is k$_3$. The reverse reaction rates are designated with negative subscripts. The overall forward (k$_f$) and reverse (k$_r$) reaction rates for the three step processes are

$$k_f = \frac{k_1}{1 + (k_{-1}/k_2) + (k_{-1}k_{-2}/k_2k_3)}$$

$$k_r = \frac{k_{-3}}{1 + (k_3/k_{-2}) + (k_2k_3/k_{-1}k_{-2})}$$

4.16

Note that the final forward and reverse reaction rates have significant differences in their parametric dependencies. Thus, the rate-limiting step for the forward and reverse multistep reactions may be different. For example, the forward reaction may be limited by formation of the reactive complex, while the reverse reaction may be limited by the reverse reaction rate [i.e. k_{-2}]. The overall rate of both k_f and k_r is generally dominated by the slowest step. When all forward reaction rates are much faster than any reverse reaction step then the denominator of the formula for k_f becomes unity and $k_f = k_1$. When one of the reverse steps is fast, for example large k_{-1}, this will lead to a large second term in the denominator if $k_2 < k_3$. In this case, the forward reaction rate becomes $k_f = k_2 (k_1/k_{-1})$. Similar analyses may be made for each of the various possible rate limitations. Various changes to the reaction conditions may supply or remove reactants or products and consequently modify any of the individual rates. In this case each rate constant must be multiplied by the appropriate reactant and product concentrations. Temperature also affects different processes in different ways. Thus, it is often possible to design a process to favor a given reaction over another. This is fortunate because we often want to select a desired event and suppress an undesirable one.

4.4.2 Nucleation

For a synthetic chemist the reaction rate is often the most crucial aspect of kinetics. Synthetic chemistry generally treats reactions in the gas or liquid phase where reactants have high mobilities and where distinct second phases are not normally formed. In semiconductor processing, reaction kinetics such as discussed above is essential to modeling techniques such as chemical vapor deposition (Chapter 12). Solid state chemistry and most materials science of solids is concerned with atomic transport as much as with reactant and product concentrations or thermodynamics. Furthermore, it is often the case that the forward reaction will not occur rapidly until some product is formed. This is normally referred to as a problem of nucleation and is similar to the classic conundrum, "which came first, the chicken or the egg?" In the remainder of this section we will examine nucleation and then turn to diffusion transport kinetics in the following section.

The phenomenon of nucleation is commonly observable in social situations, which may serve to illustrate some of the issues. Consider a party at which everyone is free to take food from a table. The food has been lavishly prepared and all of the dishes look very special. It is common that the guests will hesitate to take any of the food until someone has cut into it and spoiled the perfection of the presentation. Likewise, in a classroom students are often reluctant to ask questions. However, if someone interrupts the lecture with one question, often many more will follow. The same situation occurs in solids.

Suppose that a material exists as a single phase (say, an A + B alloy) at a high temperature. As the material cools, the phase diagram suggests that it should want to change to a two-phase mixture (of A and B phases). With all of the atoms uniformly mixed in the alloy, it would be difficult for the atoms to know what to do to form

phase-separated regions. The only chance is to rely on random events to bring a relatively large concentration of A or B atoms together locally. Such a configuration, being more stable, would tend to last relatively long. If it lasted long enough to add additional A or B atoms, then this phase would grow. Conversely, if the cluster of one type of atoms were not sufficiently stable, then it would shrink.

To further complicate the problem in a solid, nucleation of a new phase requires that a small volume of that phase be carved out of the existing materials. This creates a volume of the new phase but also a surface or interface area. If the reaction is favorable, then the energy change associated with the formation of the new volume of material should be negative (exothermic) and scales as the volume of new material created. For a spherical volume, that would be as the radius of the new phase cubed or r^3. Surfaces and interfaces are generally unfavorable (positive energy change) and have an area that rises as r^2. For an interface energy γ_s, a volume energy change ΔH, and assuming that the energies have no size dependence, the total energy of the system (i.e.: Equation 4.11) includes a term:

$$E = \pi r^2 \gamma_s + (4\pi/3) r^3 \Delta H \qquad \qquad 4.17$$

Normally, γ_s is positive and ΔH is negative for a favorable reaction with an unfavorable interface energy. For this case the total energy increases roughly as r^2, for small r, reaches a maximum, and then decreases, ultimately converging to an r^3 dependence. This situation is shown schematically in Figure 4.13.

The energy maximum occurs where

$$\frac{\partial \Delta E}{\partial r} = 2\pi\gamma_s r^* + 4\pi\Delta H r^{*2} = 0. \qquad \qquad 4.18$$

From which,

$$2\pi\gamma_s = -4\pi\Delta H r^* \qquad \qquad 4.19$$

and

$$r^* = -\frac{\gamma_s}{2\Delta H} \qquad \qquad 4.20$$

where r^* is referred to as the critical radius of a spherical nucleus. This has a positive value when γ_s is positive and ΔH is negative, as assumed above. Any addition of an atom to the cluster when $r < r_c$ has a net energy cost (the total energy increases), while beyond r^* addition of atoms is exothermic. As the surface energy increases relative to the energy gain from the reaction, the critical radius increases. The energy barrier to nucleation E^* is

$$\boldsymbol{E^*} = \frac{\pi}{12} \frac{\gamma_s^3}{\Delta \boldsymbol{H}^2} \qquad \qquad 4.21$$
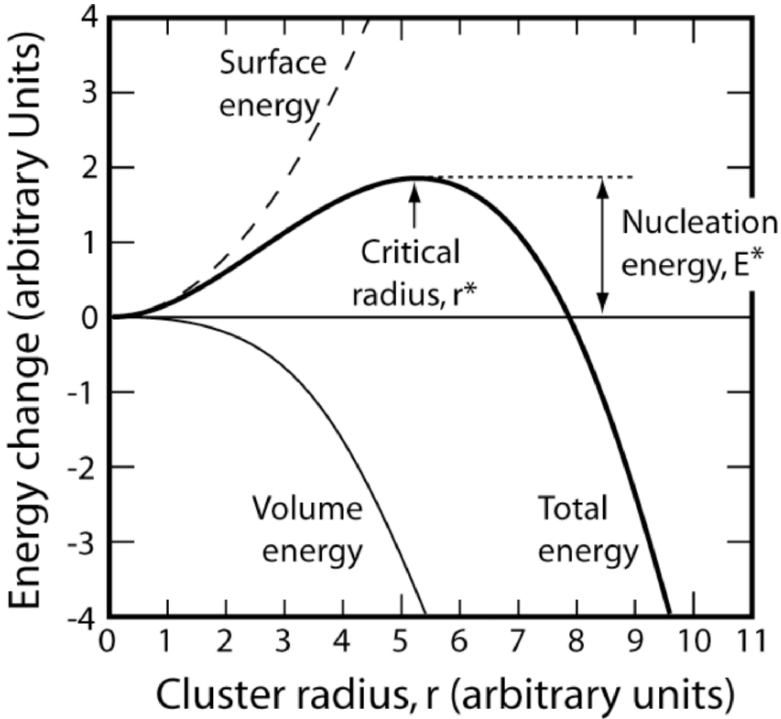
Figure 4.13: Shows the energy of a cluster of atoms during nucleation as a function of cluster size. Surface energy generally increases as r^2, while volume energy decreases as r^3. The energy of the cluster increases with addition of atoms up to a critical size, r^*, beyond which the energy decreases with atom addition. The maximum energy at r^* and is the activation energy barrier for nucleation.

for a spherical nucleus. In general, the nucleation rate, R, will depend exponentially upon the size of this barrier as

$$R = R_0 \, e^{-E^* / k_B T},$$
4.22

where R_0 is an attempt frequency. The attempt frequency will generally be energy dependent as well and may include a diffusion activation energy or some similar term. If E^* limits nucleation then point (4) in the list of considerations for reaction rates is most critical. If R_0 is the most significant limitation to nucleation rate, then point (3) in the same list dominates the overall rate. Because for a product of exponential terms the exponents may be added, the overall temperature dependence of nucleation may include a higher or lower energy than E^*. For R_0 including a diffusion rate $D = D_0 \exp(-E_D/kT)$, where E_D is the activation energy for diffusion of atoms, the nucleation rate may, for example, have an exponential dependence upon

temperature consistent with an overall energy $E = E^* + E_D$. In other words, the nucleation and diffusion barrier terms are combined in the final rate expression into a single behavior with the same general form as Equation 4.22.

Nucleation rates can present severe limitations to the speed of processes. In my own work we were coating glass slides with CdS by evaporation but were finding that for moderate exposures there was no detectable coating on the slide. Only a short time later more than one micron of CdS was found. Extrapolation of film thickness vs. deposition time data showed a delay of several minutes before growth began. This is known as a nucleation delay and was the result of a high nucleation barrier. The CdS was landing on the slide at a high rate, but would remain only a very short time in the absence of large particles of CdS to which to stick. Similar nucleation barriers are observable in many systems.

Nucleation is an essential element to most phase transformations and to growth of one material on another (heteroepitaxy). It is not an issue when the two materials are the same (homoepitaxy) or so similar as to make little or no difference. Nucleation often determines the microstructure of thin films grown from the vapor phase as well as bulk materials produced by cooling from a melt.

A final consideration is the presence of unintended or alternate processes. Often reactions or phase changes can operate along a number of different pathways. One of the oldest commonly known cases of this type (not understood in detail until the mid 20th century) was the martensitic transformation in steel. Martensite forms when a blacksmith heats a piece of medium-carbon steel to orange heat (above ~740°C) and quickly quenches it (usually in water). A slow cooling leads to a relatively soft but thermodynamically-stable two-phase material (pearlite), while the sudden quench produces a hard brittle phase known as martensite. The issue of whether pearlite or martensite forms is one of diffusion kinetics for carbon in the steel.

A second familiar example of parasitic reactions is loss of charge or loss of capacity in rechargeable batteries. In the classic case of lead-acid storage batteries, normally used in automobiles, Pb reacts with sulfate ions (SO_4^{-2}) in the sulfuric acid electrolyte solution to produce $PbSO_4$ and electrons. The electrons flow through the outside circuit to the other electrode where they allow PbO to react with the sulfuric acid (H_2SO_4) to produce $PbSO_4$ and H_2O. During charging, supplying excessive voltage can drive a competing reaction producing hydrogen but not resulting in net charge on the battery. A similar reaction in a stored battery can result in hydrogen generation and discharge of the battery without current flow through the external circuit.

Parasitic reactions also occur in electronic device fabrication. Thus, in chemical vapor deposition of AlAs one can use $Al(CH_3)_3$ (trimethyl aluminum) or $Al(C_2H_5)_3$ (triethyl aluminum) as the source gas for Al. The triethyl compound reacts with

AsH$_3$ to produce AlAs. However, the trimethyl compound can also decompose to produce AlC and hydrogen. This parasitic reaction leads to large carbon contents in the resulting films when the trimethyl compound is used.

4.4.3 Atomic transport

The common factor controlling reaction and interaction rates is atomic transport. In the vapor phase and somewhat less in liquids, transport is relatively rapid. Even so, transport often limits reactions. For example, in normal fires fuel is released from the burning material into the gas but cannot burn immediately as all of the oxygen has been used up near the burning surface by previous reactions. Combustion only begins when sufficient oxygen has mixed with the flame to allow the reaction. The flame temperature increases as the fuel-to-oxygen ratio approaches the ideal value, decreasing again as the mixture becomes fuel-poor. Most of this mixing is through convection and turbulence in the gas, but on a smaller scale, diffusion controls the mixing rate. Nonetheless, mixing is rapid enough to allow the fire to burn. Similar mixing processes occur in liquids.

In solids atomic transport is much slower. Particularly in crystalline materials the atoms have well-defined positions with strong forces holding each atom in place. Diffusion in crystalline solids is generally divided into two classes, substitutional and interstitial, depending upon whether the atom is moving primarily from lattice site to lattice site or from interstitial site to interstitial site. There are several factors that must be overcome for diffusion to occur. First, there must be a vacant site for the atom to move into. This is primarily important for substitutional diffusion as in interstitial diffusion most or all of the surrounding interstitial sites will be vacant. However, interstitial diffusion may require transfer of an atom from a lattice site to an interstitial site before diffusion will occur. There are some mechanisms for substitutional diffusion that extend the range over which one can look for a vacant site but such a site is still required to be close to the moving atom. Second, generally even for movements to the nearest-neighbor site in a lattice is it necessary for some of the atoms on surrounding lattice sites to move out of the way of the diffusing atom. This movement is normally a result of thermal vibrations of atoms in the structure and contributes to the temperature-dependence of the diffusion process. Third, as the diffusing atom moves atomic bonds are distorted or broken and reconstructed. This bond distortion contributes most of the energy barrier for interstitial diffusion. In substitutional diffusion the energy required to form vacancies is usually the largest term in the diffusion activation energy. The processes of substitutional atomic diffusion in an fcc lattice are shown schematically in Figure 4.14. Interstitial diffusion is nearly the same with the exception that the moving atom starts and ends on an interstitial site.

Because there is energy required in the diffusion process, which must be obtained from the heat of the crystal lattice, the diffusivity (related to the diffusion rate) depends exponentially on temperature following a typical Boltzmann relationship:
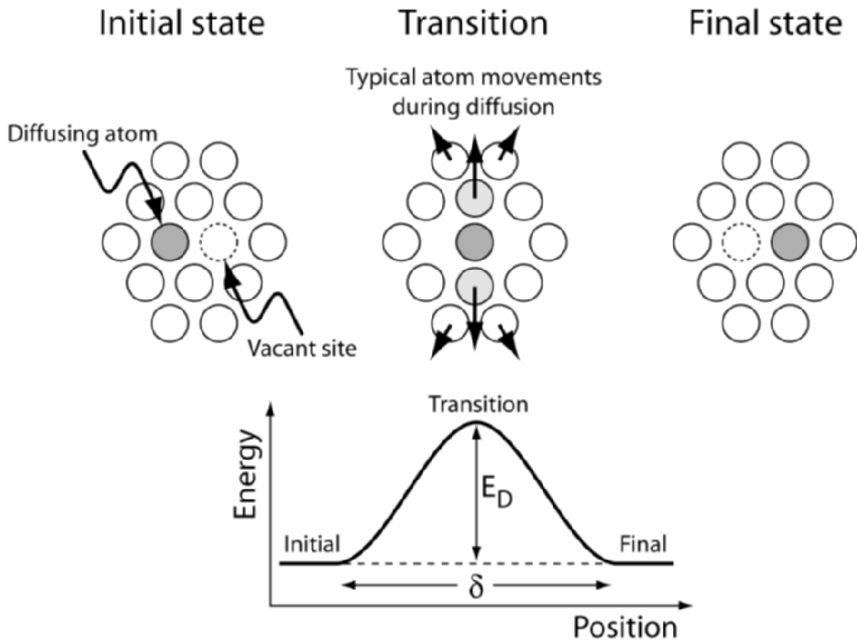
Figure 4.14: A schematic of the atomic movements involved in a typical substitutional diffusion process in a fcc lattice and the change in system energy throughout the process.

$$D = D_0 \, e^{-E_D / kT} \qquad\qquad 4.23$$

where E_D is the activation energy. The prefactor $D_0 = \nu\delta^2$ involves both the attempt frequency, ν [which is roughly the lattice vibration (phonon) frequency consistent with the lattice temperature] and the jump distance δ. In general, stronger bonds will give higher vibrational frequencies, higher energy barriers, and slower diffusion.

Diffusional mixing is usually the result of a series of uncorrelated atomic movements of the type described above. Because the movements are not related to one another, any individual atom may move in any direction. However, the net movement of atoms depends upon the concentration gradient for a given species. The result is that the flux of atoms F across a given plane perpendicular to the direction of diffusion, x, may be described by Fick's first law:

$$F = -D\frac{dC}{dx} \qquad\qquad 4.24$$

where C is the concentration of the diffusing species. This is the same relationship that appeared in Equation 3.2 and the diffusion portion of Equation 3.7 for motion of electrons in a potential gradient. One may generalize the concentration gradient in terms of a chemical potential gradient for the diffusing species and arrive exactly at Equation 3.2 with the proportionality constant related to the diffusivity of the atoms rather than to the mobility of electrons.

Kinetic limitations such as diffusion, nucleation, and reaction rates have been the basis for stability of several new classes of compounds and for the development of techniques such as rapid thermal processing. For a full description of the details and mathematics of atomic transport in solids see one of the many texts on the subject, for example, the classic text by Shewmon.

4.5 ORGANIC MOLECULES

When we discuss organic semiconductors in Chapter 9, we will be discussing organic molecules and making some use of organic chemistry concepts. It is not practical to present a summary of organic chemistry or organic synthesis here, nor is a detailed understanding of organic chemistry needed to follow the concepts of Chapter 9. However, familiarity with some nomenclature, definitions and approaches will be useful. Therefore, in this section only a few of the more important items from Freshman and organic chemistry will be considered briefly.

Organic molecules are constructed on carbon backbones. For examples of organic molecules see Figure 4.15. Most of the carbon atoms are sigma-bonded via sp^3 hybrid molecular orbitals (see Chapter 5 for more discussion of bonding). Carbon-carbon bonds define the geometry of most structural units of the molecule. Those containing exclusively carbon backbones with tetrahedral sp^3 single bonds are referred to as alkanes. Familiar members of this family include propane, butane, and so forth. Double carbon-carbon bonds (sp^2-sp^2 σ + p-p π bonds) have triangular planar molecular orbitals in their backbone bonds. Molecules containing double bonds are known as alkenes, such as ethylene and propylene.

One of the most important distinctions of alkenes from alkanes is that the single bonds can rotate while the double bonds cannot. This greatly modifies the physical properties of the molecules. Finally, alkynes include triple bonds (sp-sp σ + 2(p-p) π bonds). These are similar in many respects to the double bonded alkenes and are rigid and linear. All single and multiple carbon-carbon bonds are purely covalent, strong, and have well-defined bond angles and lengths. Interactions among organic molecules are much weaker than within the molecule and are dominated by either by van der Waals bonds or hydrogen bonds. Van der Waals bonds are important between molecules having no charged species (nonpolar) such as polyethylene. When charged subunits are present on the chain the charges may interact electro-statically to produce the stronger hydrogen bonds. The difference between polyethylene

Octane, [C_8H_{18}]

Benzene, [C_6H_6]

Octene, [C_8H_{10}]

Hydrogen

Carbon
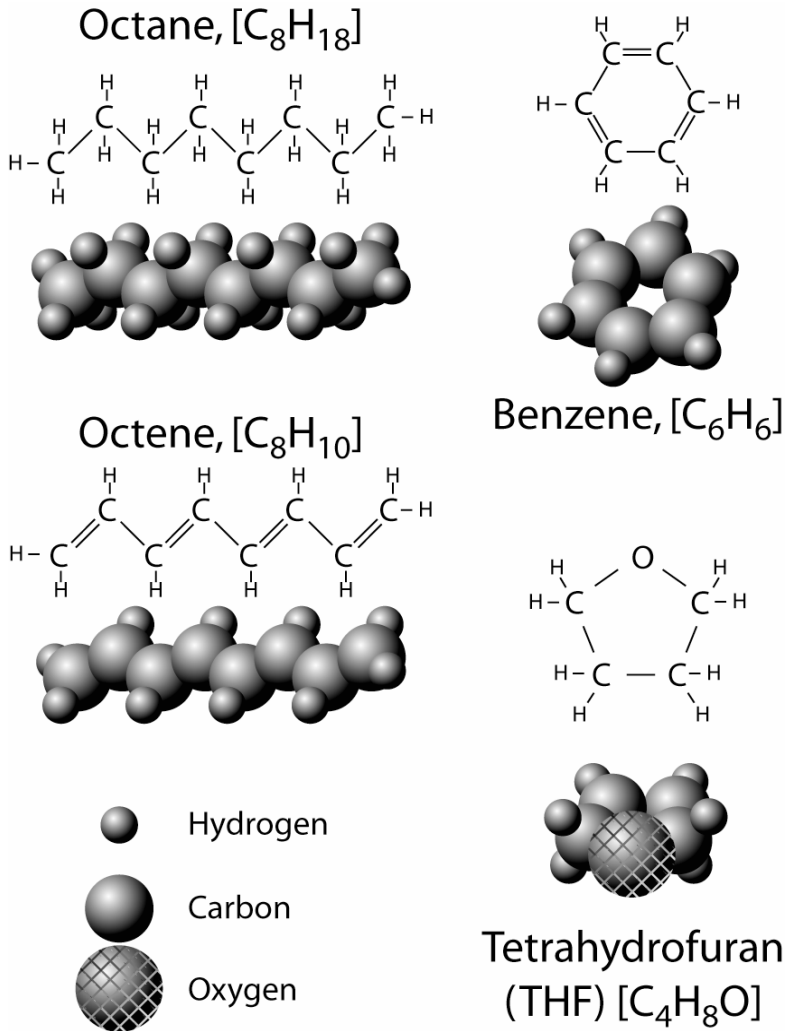
Oxygen

Tetrahydrofuran
(THF) [C_4H_8O]

Figure 4.15: Chemical formulas and sketches of the molecular structure of selected organic compounds.

and polyamide (for example nylon) is largely due to the hydrogen bonding present between charged amide units (CONH) which are absent in otherwise identical poly-ethylene.

The simplest organic molecules consist only of carbon chains with the remaining bonds terminated by hydrogen (see Figure 4.15). The properties of the materials are

easily modified, however, through replacement of a hydrogen with a larger organic group. A common example might be the difference between ethane (a gas used for cooking and heating) and ethyl alcohol, found in beverages, which differ only in the replacement of one hydrogen with an -OH group. Very effective, if time-consuming, electronic structure calculations exist by which it is possible predict some of the underlying properties of molecules with selected side groups. Therefore, it is increasingly possible to engineer a molecule to have desired properties. Modern drug design utilizes these methods extensively. In microelectronics, advanced photoresists are also developed using such tools. The challenge is often to find an effective means to synthesize the molecule, once the desired structure is predicted. The vast field of organic synthesis deals with this problem. Its methods are fascinating and elegant but far too extensive to review here. The reader is referred to the many books on organic and, for that matter, inorganic chemistry for more details.

Organic molecules are often divided into small molecules and polymers. Small molecules include most of the species that make up the familiar chemicals we use in laboratories every day. They also include more unusual materials synthesized for specific tasks. Small molecules usually dissolve readily in appropriate solvents and have properties that vary significantly with molecular weight. The differences in properties with weight make them easily distinguishable chemically and physically. As a result, they are usually separable. Thus, octane in gasoline is easily distinguished from the lighter hexane molecule. Some molecules have a variety of forms depending upon where groups or atoms are attached. This results in the "cis" or "trans" conformations and the "ortho" "meta" and "para" arrangements. These varieties are illustrated in Figure 4.16 for selected molecules. The *cis* and *trans* forms of molecules often twist along their length either in a "right handed" or "left handed" form. Virtually all biologically-active molecules twist in a "right handed" manner.

Polymers are chains of identical small molecules (monomers), strung together like beads on a string, in branched structures, or in networks. They also include copolymers that consist of more than one type of monomer in the chain. The different monomers in copolymers may be regularly or irregularly arranged. Commercial synthetic polymers such as polyethylene often have considerable variation in the lengths of the chains. This leads to variations in the properties of the material on both a microscopic and a macroscopic scale. However, the properties of individual chains do not normally vary enough to allow them to be separated chemically on a large scale basis. (Electrophoresis and related techniques do manage this separation but are slow and work with small volumes of material.) Natural polymers such as proteins synthesized by living cells, by contrast, frequently have remarkable uniformity. Most natural proteins are complex strings of selected monomer units arranged in extremely specific sequences of extremely specific lengths. This uniformity is what makes the enzymes in our bodies operate effectively.

Many polymer families have similar backbone chains but with different side groups attached to the molecule. Thus, polyethylene is a bare carbon backbone with hydrogens attached, while polystyrene has 1/4 of the hydrogens replaced with six membered phenyl rings. These side chains are often referred to as *functional groups or ligands*. An additional term used in organic chemistry is *moiety*, a part of a molecule. Thus the -CO-NH- unit, which is present in all polyamides, is a moiety. Benzene moieties are often referred to as *aromatic* groups because they determine the smell of many odorous molecules. Examples of molecules with different functional side chains and ligands are shown in Figure 4.17. This figure also shows equivalent structural representations for two of the molecules to illustrate organic chemistry short hand.
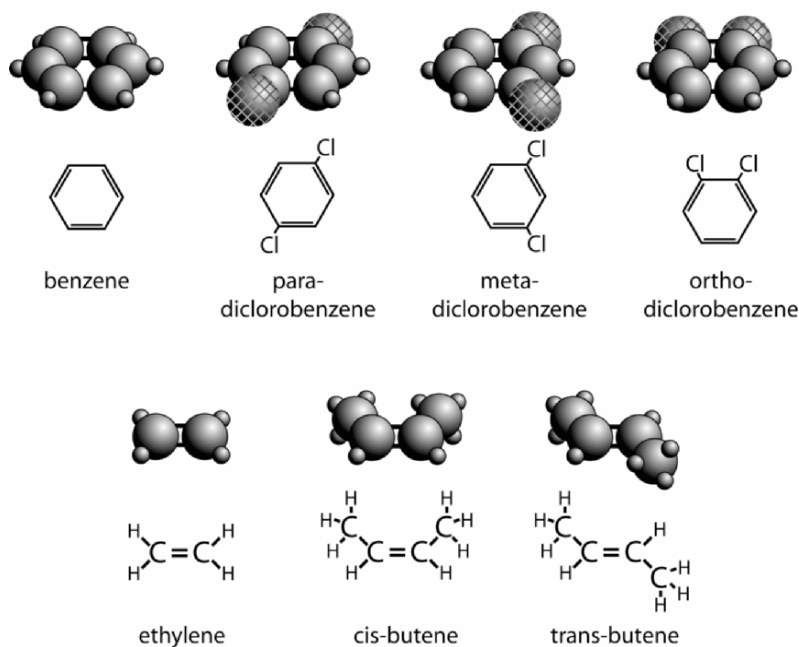

| benzene | para-
diclorobenzene | meta-
diclorobenzene | ortho-
diclorobenzene |

| ethylene | cis-butene | trans-butene |

Figure 4.16: Shows conformations of several organic molecules to illustrate the *para*, *meta*, *ortho*, *cis*, and *trans* organization of moieties on the central structure of the molecule. If one takes the $H_3CHC=C$ portion of the *cis-* and *trans*-butene molecules as a fixed reference frame one can see that this causes the molecule to turn right or left down its length. The *para*, *meta*, and *ortho* arrangements of functional groups such as Cl determines the distribution of electric charge on the benzene ring structure. This, in turn, modifies the chemical behavior of the molecules above and beyond changes directly related to the organization of the functional groups themselves.
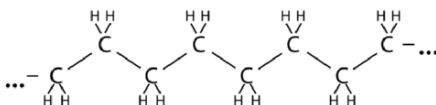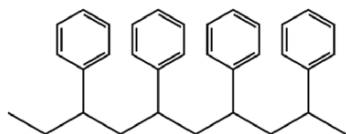
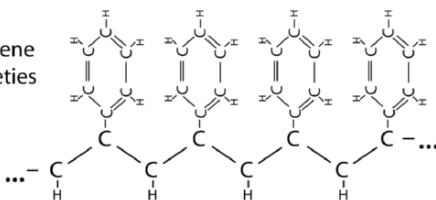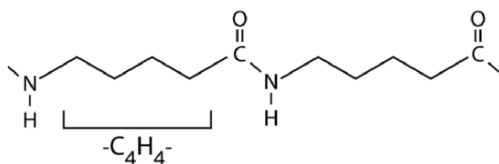*Figure 4.17:* Chemical formulas for three polymers. Structures are shown in three forms. The basic chemical formula is shown after the common name, the structure including all atoms is shown for two of the structures, and the common chemical shorthand drawing is given for each structure.

Because showing all of the atoms in a molecule produces a very busy drawing that is hard to interpret and because moieties often are turned with respect to the drawing, it is simpler to abbreviate some aspects of the structure. As shown in Figure 4.17, the simplified structures do not include an indication of hydrogen. Any left-over bond to a C atom not specifically indicated is presumed to be a bond to a hydrogen (not shown). A dangling bond is generally indicated by a dot indicating an unpaired electron or two dots if the electrons are paired (see Chapter 9 for examples). C atoms are also not shown explicitly on this diagram. C atoms are presumed to be present at vertices where lines indicating bonds meet if no other atom is shown. Non-C atoms are shown explicitly.

Aspects of functional groups which change the behavior of a molecule include 1) polarity or electronic charge, 2) size or shape, 3) reactivity, etc… Polarity is essential to the interaction of one molecule with another, for example through the formation of hydrogen bonds. This is also evident in solubility of common materials. Polar materials generally dissolve in polar materials and non-polar materials dissolve in non-polar materials, but the two do not usually mix well. Water and alcohols are examples of polar materials, while methane, hexane, and most oils are non-polar. Polymers or small molecules may include both polar and non-polar segments. This controls how they react and interact. In such cases the polar segment dissolves in a polar solvent while the non-polar parts prefer to clump together. This is how micelles form and is also the basis for the cell wall in biological systems.

Polymers are formed by reactions joining small molecules together to create the final structure. When the small molecules have two reactive groups, reactions joining them result in straight chains. Straight chain polymer molecules may be relatively rigid or relatively flexible, depending upon the flexibility of individual bonds in the chain. Ring structures are always inflexible units. Monomers with more than two reactive sites produce networks rather than chains. In this case the entire solid becomes effectively one huge molecule. The extent of interconnection in a polymer results in dramatic changes in its properties and divides polymers into two groups, thermoplastic and thermosetting.

Thermoplastic polymers consist of individual molecules, usually of variable length but relatively uniform chemistry, entangled with each other but not directly bound together by covalent chemical bonds. Weak interchain bonding may further organize the polymer chains into sheets or small crystallites if the polymer has few side groups and is very regular along its length. Because they consist of a jumble of individual molecules, thermoplastic polymers may melt if heated or may dissolve in solvents. Examples of thermoplastic polymers are polyethylene and polystyrene.

Thermosetting polymers consist of networks of monomers. It is common to make thermosetting materials from thermoplastic materials by adding small molecules, which will link the thermoplastic polymer chains together with covalent bonds to form a network. Epoxy is a common example of such a material. The epoxy resin is a thermoplastic polymer with reactive side chains. The hardener is a cross-linking molecule that links the resin molecules together. Because of the strength of the individual C-C bonds, thermosetting materials cannot be melted. Because the entire material is one huge covalently-bonded molecule, they do not dissolve in solvents, although solvents often will intercalate into the network and cause the material to swell. Thermosetting materials tend to be hard and brittle because it is not straight-forward for them to deform.

4.6 APPLICATIONS

4.6.1 A basis for phase transformations

To apply the formulas in Section 4.2 in more detail, one can, for example, calculate the free energy of a mixture of materials, from which the binary phase diagram may be obtained. Consider a mixture of two elements, A and B with composition x. The composition of this mixture would then be designated A_xB_{1-x}. A and B might be Pb and Sn in the case of Figure 4.8 but we will consider them generically here. Under the regular solution model these two elements can form three types of bonds, A-A, B-B, and A-B with distinct energies. It is assumed in this model that individual pairs have well defined energies and that there is no distinction among A-B bonds in which the A atom is also bonded to other A or B atoms. In real materials, especially semiconductors, the other atoms surrounding a given atom have a strong effect. The presence of significant interfacial energies also suggests that the pairwise bond energy assumption is highly approximate. Nonetheless, we will use it here. These considerations will be seen to be important when we discuss semiconductor band gap bowing and miscibility criteria in Chapter 6.

The difference in energy of two A-B bonds (energy=$2\omega_{AB}$) relative to one A-A (ω_{AA}) and one B-B (ω_{BB}) bond, is $\xi = 2\omega_{AB} - \omega_{AA} - \omega_{BB}$. ξ defines the energy gain or loss upon mixing of the elements. In a random alloy the probability of A-B bonds is the fraction of A atoms multiplied by the fraction of B atoms or x(1-x). The average bond energy, Ξ, of the resulting material is thus:

$$\Xi = 2x(1-x)\xi \qquad\qquad 4.25$$

This is also the energy of mixing of A and B atoms as it is the average change in energy per bond of replacing separated A and B atoms with a random mixture. Thus, if positive the mixing is endothermic and if negative the mixing is exothermic. The latter conclusions are only true at zero Kelvin because at elevated temperatures the entropy of mixing also enters the problem.

The primary term in the entropy of mixing is the configurational entropy, ΔS_{mix}. This is calculated by analysis of the number of combinations of N_A A and N_B B atoms which can be placed on the total number of lattice sites N_A+N_B in the solid. The number of configurations is:

$$\Omega = \frac{(N_A + N_B)!}{N_A! \quad N_B!}, \qquad\qquad 4.26$$

from which the entropy is $\Delta S_{mix}=k\ln\Omega$. The logarithm of a factorial can be estimated with Stirling's approximation, $\ln(N!)\sim N\ln(N)$ from which:

$$\frac{\Delta S_{mix}}{k_B} \approx (N_A + N_B)\ln(N_A + N_B) - N_A \ln(N_A) - N_B \ln(N_B). \qquad 4.27$$

Taking N=N_A+N_B and given that the difference of logarithms is the logarithm of the ratio of their arguments one may show that:

$$\frac{-\Delta S_{mix}}{Nk_B} = x\ln(x) - (1-x)\ln(1-x) \qquad 4.28$$

where we have also used the fact that N_A/N=x and N_B/N=(1-x). The total free energy change upon mixing per lattice site ($\Delta F/N$) can be shown to be the negative of the sum of Equation 4.25 and T multiplied by Equation 4.28. In other words:

$$\frac{\Delta F}{N} = k_B T\left[x\ln x + (1-x)\ln(1-x)\right] - \frac{x(1-x)\xi}{2}. \qquad 4.29$$

One can then plot F as a function of x for several values of T given a value for ξ. An example of such a calculation is given in Figure 4.18.

When the energy change per lattice site is positive the system is unstable. When negative the system will choose the lowest energy composition possible. When there is only one minimum in $\Delta F/N$ then there is an energy penalty to decomposing the mixture into separate phases. However, when two minima occur there is a reason to change the material into separate domains with each of the two low-energy compositions. In short, this favors the phase transformation. Note that the calculation provides a basis for understanding compositional instabilities and hence a process called spinodal decomposition.

Spinodal decomposition is phase separation resulting from fluctuations in composition that become thermodynamically stable up to some magnitude as the temperature decreases. The scale of stable fluctuations increases as the temperature decreases. This process does not produce a clear transition from a single phase to a two phase mixture in the microstructure because the first stable composition fluctuations have a very small magnitude. Thus, at least at first, the local differences in the separated phases are tiny. As temperature is lowered and the separation proceeds the separate phases become increasingly distinct. In many phase separation processes the initial small composition fluctuation regime is masked by the presence of a soluble liquid phase. Thus, for example, the spinodal decomposition marked by the solid line in Figure 4.18c is often cut off at the top as indicated by the dashed lines in the figure. Thus, by the time the phase separation begins the stable compositions are already far separated.

Full binary phase diagrams generally include other issues such as melting or boiling. Thus a complete treatment of the phase relationships in a mixture will involve additional calculations to estimate the energy of the other phases. The stable compositions and phases are those with the lowest overall energy. A consequence of
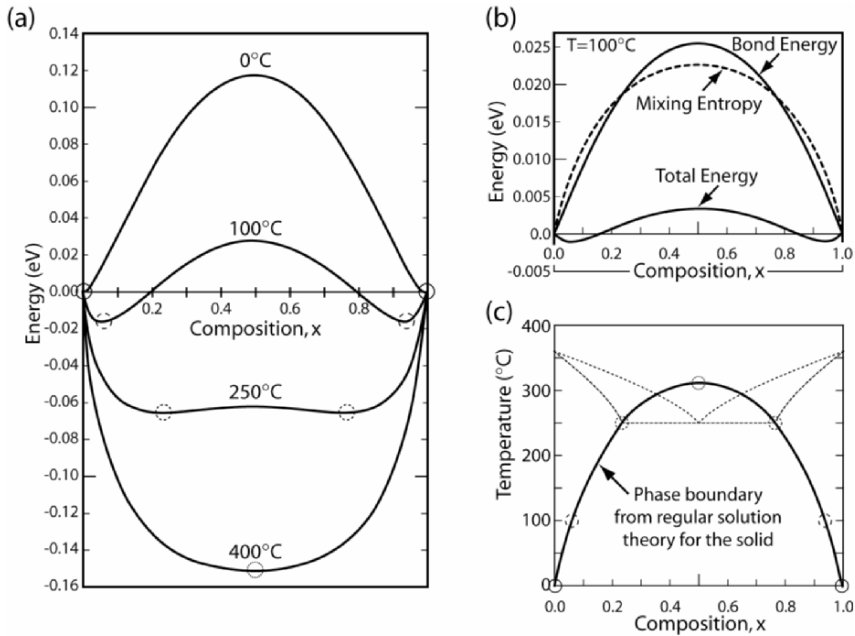
Figure 4.18: Free energy as calculated from Equations 4.25-4.29. (a) shows the free energy of mixing for several temperatures. Negative values indicate favorable mixing, while positive values indicate a preference for demixing. The more negative the energy the more favorable the mixing. (b) gives the two contributions to the energy, bond enthalpy and heat (temperature x entropy), whose difference gives the curves in (a). Notice in particular that the entropy dependence upon x is significantly more square than the energy dependence. (c) the locus of energy minima from (a) [circles] as a function of temperature determine the boundary of the stable phase regions and the compositions which are stable in the two-phase region. This locus determines the shape of the phase boundaries in the solid phase region of a binary phase diagram. To give an idea of how the rest of the phase diagram enters, the dashed curves are added. Calculation of these requires a separate determination of the energy of the various liquid phases.

including the liquid phase in the calculation would be the potential appearance of other phase boundaries such as are shown in the phase diagram in Figure 4.19.

4.6.2 Silicon crystal fabrication

Several of the concepts in this chapter are illustrated in the refining of Si from natural quartz or silica sand to electronic-grade Si and the subsequent growth of single crystals by the Czochralski process. Natural silica sand and quartz are nearly pure SiO_2. Unfortunately, getting rid of the oxygen in SiO_2 often degrades this initial purity dramatically. The reduction of SiO_2 is accomplished by mixing with a carbon source, normally coal, and other materials and heating the mixture to a high temperature

(chosen to melt the component species, provide a rapid reaction, and favor the reduction of Si). The basic reduction process is shown in Figure 4.20.
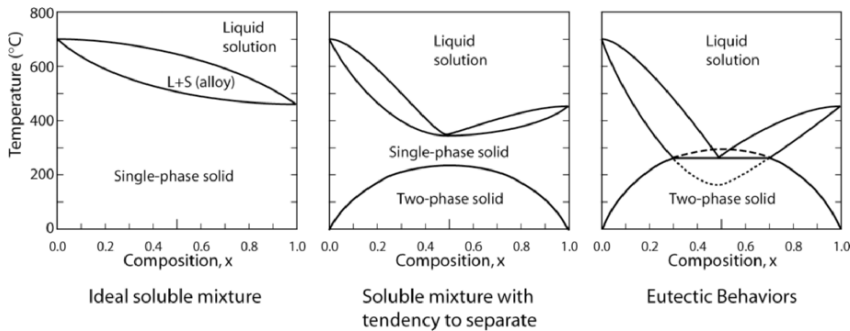

Figure 4.19: In a full calculation of a phase diagram one can include the liquid phase. An ideal solution produces a liquid-solid phase diagram as on the left side of the figure. As the bond energy difference from Equation 4.25 becomes more significant the solid tends to prefer to separate into two phases. This has consequences for both the low-temperature behavior described by Equation 4.29 and Figure 4.18 and for the liquid-solid relations. The result is the middle figure. A further increase in the tendency to phase separate leads to the behavior on the right and to a binary eutectic phase diagram as in Figure 4.8.

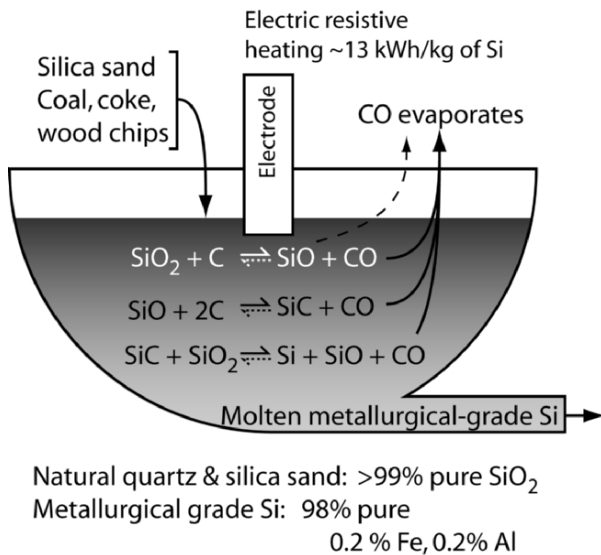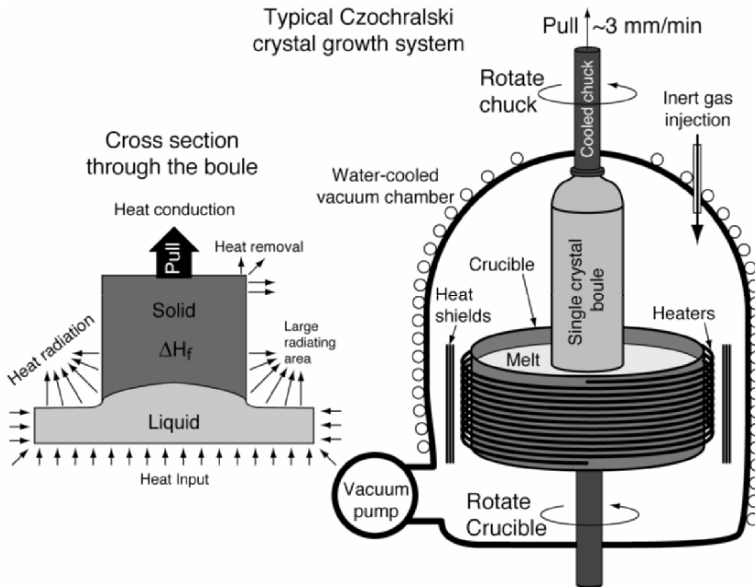*Figure 4.20:* A schematic of the process for refining SiO_2 into metallurgical-grade Si using reduction by C. The product CO is removed by evaporation from the surface and is swept away in the gas. Because SiO has a moderate vapor pressure at these temperatures, some SiO also leaves the surface and condenses on cool surfaces nearby.

Oxygen reacts preferentially with C rather than Si making the reaction $SiO_2 + 2C \rightarrow Si + 2CO$ exothermic. [Note that CO is favored over CO_2 at high temperatures.] The removal of CO from the gas makes the process effectively irreversible (from Equations 4.14 and 4.16.) Thus, the Si is efficiently reduced, leaving only moderate O and C contamination. Unfortunately the C sources, such as coal, are impure and introduce contaminants, notably iron. These must be removed by subsequent processing. The result of the reduction is metallurgical grade Si.

The behavior described by Equation 4.14 is put to practical use as well in the purification of Si based on the reaction: $Si + 3HCl \rightarrow SiHCl_3 + H_2$. This reaction is exothermic and runs forward rapidly at 300°C in the presence of high concentrations of HCl gas and with removal of the gaseous reaction products. The tricholorosilane produced is easily separated from the unreacted HCl by condensation on a cool surface. It can then be fractionally distilled for purification. The reverse endothermic reaction forming Si and HCl will occur at higher temperature (~900°C) in the presence of hydrogen. This takes advantage of providing excess H_2 reactant and sweeping away product HCl in the unused H_2 to prevent the exothermic reaction. The trichlorosilane process was the standard method for purification of Si for many years. It is very straightforward but energy intensive and somewhat slow. A more efficient and speedy alternative has been developed based on silane (SiH_4) formation and decomposition. Both the Si reduction and purification processes are enhanced through the law of mass action. By removing reaction products the reverse reaction is slowed, thereby effectively increasing the desired forward reaction rate.

The final step in the production wafer fabrication process is making single-crystals from polycrystalline electronic-grade Si. The Czochralski process, shown schematically in Figure 4.21, normally achieves this. A typical Czochralski crystal growing system and the resulting boule are shown in Figure 4.22. In this method a small (4-6 mm diameter) "seed" single crystal of the material to be grown (Si, for example) is mounted on a rotatable and movable chuck in a desired orientation. As the crystal grows, this initial orientation determines the orientation of the resulting boule (bulk cylindrical crystal). The seed crystal is lowered gradually into a pot of liquid Si until it barely touches the surface. To grow the crystal the seed is then slowly (a few mm min^{-1}) withdrawn from the melt. Pulling the crystal out of the melt faster reduces the boule diameter while slower pulling increases the diameter. Changing the melt temperature also affects size. The boule diameter is measured dynamically through ports in the growth chamber and the pull speed is adjusted to achieve a given size of crystal. Rotation of the seed crystal and melt is needed to control the uniformity of the melt. This is critical to quality of the resulting crystal.

Inevitably a cold seed crystal put into contact with hot liquid experiences a strong thermal shock. This produces dislocations in the seed that must be prevented from propagating into the resulting boule. To accomplish this the growing crystal diameter is reduced to 3-4 mm and then enlarged again to the desired size (Figure 4.22). The entire weight (>100 kg) of the final boule is supported on this tiny neck!

Typical Czochralski
crystal growth system

Pull ↑~3 mm/min

Rotate chuck

Inert gas injection

Cross section
through the boule

Water-cooled
vacuum chamber

Heat conduction

Heat removal

Pull

Solid

ΔH_f

Heat radiation

Liquid

Large
radiating
area

Heat Input

Crucible

Heat shields

Melt

Cooled chuck

Single crystal boule

Heaters

Vacuum pump

Rotate Crucible

Figure 4.21: A schematic diagram of the Czochralski crystal growth process. A seed crystal mounted on a cooled chuck is lowered into a molten bath, which will solidify to start crystal growth. This seed is then withdrawn slowly and new crystal grows on the existing seed. If the seed is a single crystal then with proper care the boule will also be a single crystal with zero dislocations. The diameter of the boule is adjusted by control of the shape of the meniscus of the liquid and the shape of the solidification boundary.

One of the essential issues in growth of Si single crystals is temperature management. The seed crystal is cooled by radiation to the cold walls of the vacuum chamber in which the growth occurs. At the same time, heat is supplied to the crucible containing the melt. Finally, the latent heat of freezing, ΔH_f, provides an additional source of heat located, inconveniently, at the liquid-solid interface. A temperature gradient from the seed into the melt is established by these heat sources and sinks and the solid forms where the material passes through the melting temperature. If the crystal is growing uniformly the latent heat is generated equally across the solid-liquid interface. However, because the major heat sink is radiation from the surface of the melt and the boule, heat escapes most rapidly from the edges of the boule. In the center only conduction up the boule is available to remove the heat. Hence, this is the hottest region. A heat transfer model accounting for the various sources and sinks shows that the lower surface of the boule curves upward toward the center of the crystal, as shown in Figure 4.21.
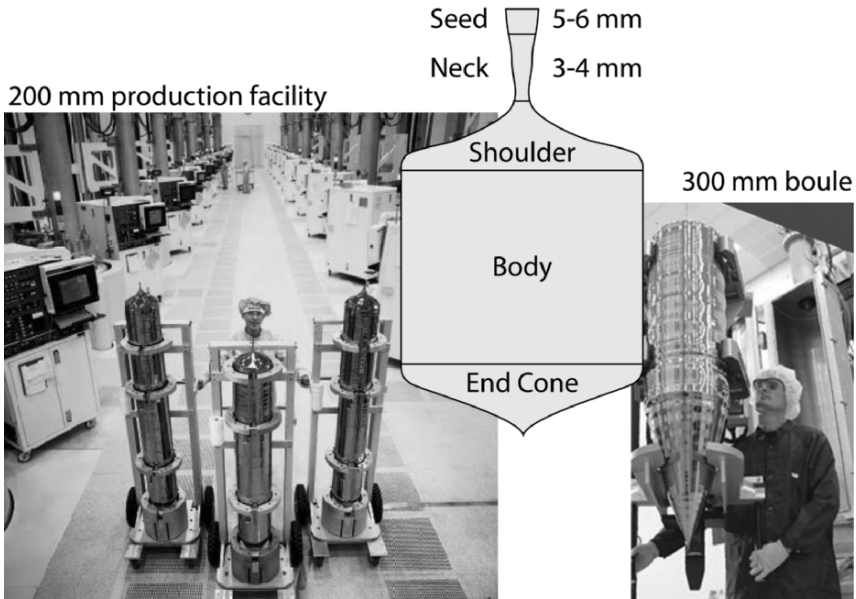
Figure 4.22: Photographs (left) of a 200-mm Si Czochralski growth facility, and a 300 mm Si boule (right) courtesy of MEMC Electronic Materials, Inc; St. Peters MO. The overlying schematic shows the regions of the boule.

The shape of the freezing surface is crucial to good crystal growth, as is control of convection in the melt. In this regard, the latent heat of solidification is particularly important as it can enhance convection. If freezing is nonuniform, local bursts of heat can be generated. When the boule surface is tilted, as is the case if it is significantly concave in the melt, a hot spot may be carried laterally across the boule surface and can cause melting of an adjacent area. These heating and cooling bursts produce locally high crystal growth rates, which can result in defects in the solid, as well as variations in doping. Convection currents in the melt are reduced in current Czochralski pullers by application of a large magnetic field (using superconducting magnets in some cases). The magnetic field interacts with moving atoms in the melt and acts as a brake, slowing convection significantly.

A modern Czochralski crystal growth system producing 300 mm-diameter 1200 mm long Si boules from a charge of Si in the crucible might typically stand 10 meters high, consume ~1 MW of electric power and over 760 L of cooling water and 400 L of Ar per minute. Roughly 2/3 of the electric power goes to energizing the magnetic convection-damping field (suggesting the motivation for a superconducting magnet).

The crucible generally holds ~250 kg of electronic-grade Si with both the crucible and boule rotated at up to 30 rpm. The crystal is grown at up to 8 mm min^{-1}. Thus, to grow a 1 m long boule would require ~1.5 hours to grow, consuming ~1500 kWh of power, 68000 L of water, and 36,000 L of Ar.

One of the observations about Czochralski growth is that the concentration of impurities in the solid is generally much different (and fortunately often much lower) than in the liquid. The ratio of the concentration in the solid to the concentration in the liquid is referred to as the segregation ratio. Typical values for segregation ratios of selected impurities are given in Table 4.1. Some of the more striking values are for O (which actually prefers to be in the solid phase) and Fe, which is in equilibrium when the liquid contains one hundred twenty five thousand times more Fe than does the solid. The latter is particularly fortunate as Fe is a major problem for semiconductor devices. Detectable degradations in performance may be found at the part-per-billion Fe level. Consequently, effective methods, such as those discussed above, to remove Fe are particularly important.

Table 4.1: **Segregation Ratios of Impurities in Silicon**

Impurity	C_{solid}/C_{liquid}	Impurity	C_{solid}/C_{liquid}
B	0.8	C	0.07
Al	0.002	Fe	8×10^{-6}
P	0.35	O	1.25
As	0.3		

The values in the table show why B would be better than Al for intentionally doping bulk Si grown by the Czochralski method. First, one would need 400 times more Al in the melt than B to achieve the same bulk doping level in the boule. Second, as the melt is used up in forming the solid the number of Al atoms in liquid is relatively unchanged while the Si is consumed. Thus the Al concentration in both the liquid and solid rises increasingly rapidly as the crystal grows. By contrast, because B is removed from the liquid at nearly the same rate as Si, the concentration in the melt and hence in the solid remains nearly constant.

To see why segregation might be expected, consider the composition and temperature of the liquid near the solid interface. This may be guessed based on the phase diagram of silicon and the impurity. Suppose the region of a generic binary eutectic phase diagram near the Si melting point is as shown in Figure 4.23. The "solidus" line marks the composition at a given temperature below which all of the material is solid, while above the "liquidus" line the system contains only liquid. Between the two lines the system is a mixture of solid and liquid phases.
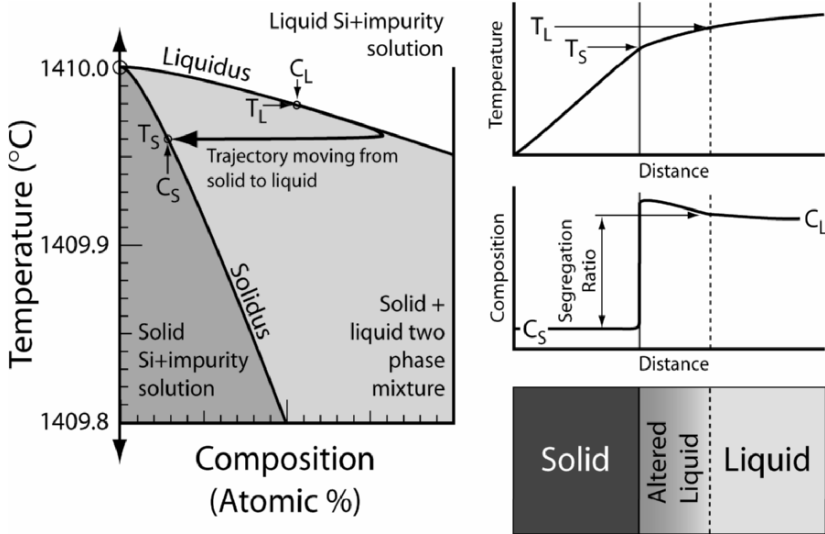
Figure 4.23: Schematic diagram of a typical binary eutectic phase diagram between Si and an impurity near the Si melting point. Insets at the right show the temperature and composition of the system across the solid/liquid interface, as well as a schematic of the interface. The binary phase diagram includes an indication of the compositions of the solid and liquid during growth and the change in both variables through the interface region. As shown, the composition of the liquid is enriched in impurity near the interface due to rejection of extra impurity from the solid during growth. This excess impurity must diffuse into the liquid and results in the composition change shown. Note that the observed segregation ratio will generally be lower during growth than the equilibrium value due to this diffusion rate limitation.

For any temperature, the segregation ratios given in Table 4.1 are the ratio of the composition of the solid to that of the liquid, given by the solidus and liquidus lines, respectively. However, near the solid-liquid interface there are two variables that could change across the interface, composition and temperature. Very close to the solid-liquid interface the temperature is nearly constant at the local freezing point for the local liquid composition. This corresponds to a horizontal line on the phase diagram, shown as the lower line on the trajectory through the interface on Figure 4.23. Because heat is generally transmitted more slowly in the solid, the temperature gradient there is likely to be larger. However, the impurity is also being rejected into the liquid as the solid grows due to segregation. Therefore, the liquid near the interface has an enhanced level of impurity determined by how quickly those atoms can diffuse away. At steady state a concentration distribution, as shown in Figure 4.23, will be established. A changing concentration of impurity in the liquid with

position will also correspond to a change (generally a rise) in liquidus temperature. As a result, both temperature and composition will change across the solid-liquid interface region, as shown schematically by the trajectory in Figure 4.23. Therefore, the temperature gradient in the altered liquid layer is connected to the composition gradient there to some extent. For infinitely slow growth where diffusion in the liquid is very fast, the segregation ratio would be the ratio of solid to liquid composition at the solid-liquid interface. For faster growth the segregation ratio is effectively reduced due to the development of a concentration gradient in the liquid.

Much of this text is concerned with semiconductor alloys, particularly of the group III-V semiconductor compounds. These materials are not produced on Si substrates, but rather rely on GaAs, InP or other single crystals. Although boules of these materials are grown by the Czochralski process, there is a catch – the group V element in the case of III-V compounds is very volatile at the crystal growth temperature and would evaporate into the gas stream if something were not done to prevent its escape. One solution is the "liquid-encapsulated" Czochralski (LEC) method. In this process liquid boric oxide (B_2O_3) floats on top of the melt used to grow the boule (InP for example). Because it is insoluble in the boule material and vice versa the interface is stable. The B_2O_3 prevents melt material from dissolving in it and escaping to the surface. Because it is stable at the growth temperature the boric oxide itself will not decompose or evaporate. Finally, because it is stable with respect to GaAs, InP, or other materials being grown it will not dissolve or allow As or P to pass through it. Thus the boric oxide layer forms a barrier preventing escape of the volatile element in the compound semiconductor. This process has been used to produce compound semiconductor wafers up to 150 mm in diameter.

An alternate method is to allow the group V element to evaporate into the growth environment and to simply raise its pressure in the crystal puller until it is condensing at the same rate that it evaporates. The problem with this approach is that it requires a very high pressure of (somewhat dangerous) group V element in the puller and that the puller be designed to withstand this pressure safely. The major advantage of this method is that modest B and O contamination of the growing wafer does not occur. This permits lower ultimate doping levels in the wafers.

4.6.3. Rapid thermal processing

As a final example of materials science concepts in action, we briefly consider rapid thermal processing, used extensively in microelectronics manufacturing today. This technique is based on the fact that most processes from diffusion to reaction rates to nucleation in thin films have kinetic rate constants that depend exponentially on inverse temperature, as in Equation 4.15. Rapid thermal processing takes advantage of the fact that for two competing processes with different values of the activation energy E, the higher activation energy process is strongly favored at high temperature, while it is strongly suppressed relative to the other at low temperature. Thus, by adjusting the temperature one can adjust the relative rates of two processes.
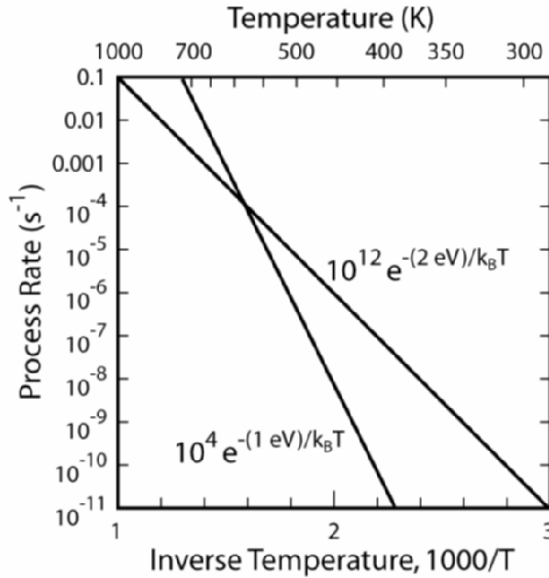
Figure 4.24: Rates for two hypothetical processes, one with an activation energy of 2 eV and a prefactor of 10^{12} sec^{-1} and a second with an activation energy of 1 eV and a prefactor of 10^4 sec^{-1}. The high activation energy process is favored by short times at high temperatures (left side of the graph) while the low activation energy process is dominant for long times at low temperatures (right).

It turns out that in many cases in microelectronics fabrication, the activation energy of a desired process is higher than for an undesired processes, as shown in Figure 4.24. Furthermore, in many processes, such as diffusion, we must only keep constant the product of diffusion rate and time (a product directly proportional to the square of the distance an atom can diffuse in that time). Therefore, by increasing the process temperature one can decrease the annealing time at that temperature. Hence the name, rapid thermal processing. While the desired process rate x time product is kept constant at the increased temperature, the undesired process rate has increased but to a lesser extent. Thus, when the time at elevated temperature is reduced, the undesired process has a relatively lower rate x time product and is suppressed relative to the desired process.

Rapid thermal processing can, in general, be used to suppress any low-activation energy process in favor of any high activation-energy process by raising the process temperature and decreasing the time. The reverse could be achieved with a low-temperature long-time heat treatment if a desired process had a very low activation energy.

4.7 SUMMARY POINTS

- All crystals form in one of 14 Bravais lattices.
- Planes and directions in these lattices are indexed with three-integers in cubic systems.
- Each lattice point on the Bravais lattice may have more than one atom associated with it.
- A "primitive cell" also may be defined containing exactly one formula unit of atoms of the compound.
- Hexagonal close packed and face-centered cubic crystals are both "close-packed" lattices having ABAB or ABCABC stacking sequences for atom planes.
- A reciprocal lattice may be defined based on Equation 4.2 and the basis vectors for the real-space lattice.
- Diffraction occurs when the change in wave vector in the diffraction event is a reciprocal lattice translation vector.
- A phase is a chemically and structurally homogeneous material.
- Thermodynamic stability requires that a phase or set of phases have a lower energy than any other set of phases that could be constructed from the same set of atoms under given conditions. This is equivalent to the set of atoms having their lowest average chemical potential.
- Entropy is the Boltzmann constant multiplied by the logarithm of the number of states accessible to the system. In reversible processes an isolated system exhibits no change of entropy while in irreversible processes the entropy must always increase. Entropy contributions to the energy of systems become increasingly important at higher temperatures.
- The number of degrees of freedom of a system is the number of constituents less the number of phases plus two. Single constituent systems have pressure-temperature phase diagrams such as that shown in Figure 4.7, binary phase diagrams are as in Figures 4.8 and 4.9, and ternary phase diagrams are as in Figures 4.10 and 4.11. Other phase diagrams also exist.
- The phases present in a multiphase region of a phase diagram are determined by the stable phases in surrounding regions of the diagram at a given temperature.
- Phase diagrams may be calculated from basic thermodynamics describing the interactions among atoms and the entropy associated with various arrangements of atoms.
- The less the atoms like to mix the more likely they are to form eutectic-type phase diagrams rather than showing large solubilities. This is determined largely by the relative energy of the bonds between different atoms compared to the bonds between like atoms.
- The law of mass action shows that the rate of a reaction depends upon a ratio of the reactants to the products multiplied by a rate constant. This rate constant

depends exponentially on temperature and an activation energy and linearly on a prefactor.

- Multistep reactions may be limited by any of the rates of an intermediate reaction or by a combination of all rates.
- Any phase change creating a new interface is likely to require nucleation of the new phases. Nucleation is generally associated with an energy barrier and may be the limiting step in a reaction or phase transformation.
- Nucleation energy barriers depend upon the surface and volume energies associated with the reaction. A critical radius for the nucleus exists beyond which the new phase may grow spontaneously.
- Phase transformations often require transportation of atoms resulting from diffusion.
- The flux of transported species depends upon their concentration gradient and diffusivity.
- The diffusivity depends exponentially on temperature and an activation energy and linearly on a prefactor, usually related to a lattice vibrational frequency and the square of the jump distance over which the particle moves in a single average event.
- Organic molecules have covalently-bonded backbones with single (flexible), double or triple (rigid) bonds along their backbones.
- Interactions between chains are through weak Van der Waals or hydrogen bonds. The latter require charged subunits on the molecule.
- Properties of organic molecules are determined by the nature of functional groups, ligands, or moieties within the molecule, and by the arrangements thereof.
- Polymers are chains (thermoplastic) or networks (thermosetting) of small monomer molecules.
- Electronic grade Si is produced by reduction by C, followed by a purification step.

4.8 HOMEWORK

1. Derive Equation 4.21 for the critical nucleus of a cluster based on Equations 4.17 and 4.20.

2. Derive Equation 4.28 from Equation 4.27.

3. The basis vectors for a face-centered cubic real-space lattice, $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ are
    $$\frac{a}{2}(1,1,0), \quad \frac{a}{2}(1,0,1), \quad \frac{a}{2}(0,1,1)$$
    calculate the reciprocal lattice basis vectors $\mathbf{b}_1$, $\mathbf{b}_2$, and $\mathbf{b}_3$.
    Determine the lattice structure of the reciprocal lattice.

4. Light a match or candle and observe the resulting flame. Make a sketch of the flame showing the match stick or candle wick and the various visible parts of the flame. Indicate approximately on your drawing:

    a) the fuel rich region
    b) the region close to the ideal reactant mixture
    c) the fuel-poor region
    d) where the fuel is coming from
   e) where the oxygen is coming from.

5. Suppose we are attempting to nucleate a solid-phase reaction between a metal and a silicon substrate to produce a silicide intermetallic compound and carry the reaction to completion. Further suppose that we have a dopant in the solid which we do not wish to have diffusing as the silicide reaction takes place. Suppose the two processes both have rates defined by Equation 4.26. Given the kinetic parameters for the reactions as follows:

	k_0 prefactor	Activation Energy, E
Silicide reaction:	1×10^{12} cm^2 s^{-1}	3.5 eV
Dopant diffusion:	4.3×10^3 cm^2 s^{-1}	2.3 eV

    The maximum temperature at which the silicide reaction can take place is 850°C and the minimum temperature is 450°C. If the product of rate and time for the desired process must be 2×10^{-7} cm^2, calculate:

    a) the time necessary for the desired reaction process at 450°C
    b) the time necessary for the desired reaction process at 850°C
    c) the time-temperature products for the undesired diffusion at these temperatures.
    d) the ratios of the two process rates at the two temperatures

6. Consider the phase diagram of an impurity in Si near the Si melting point:
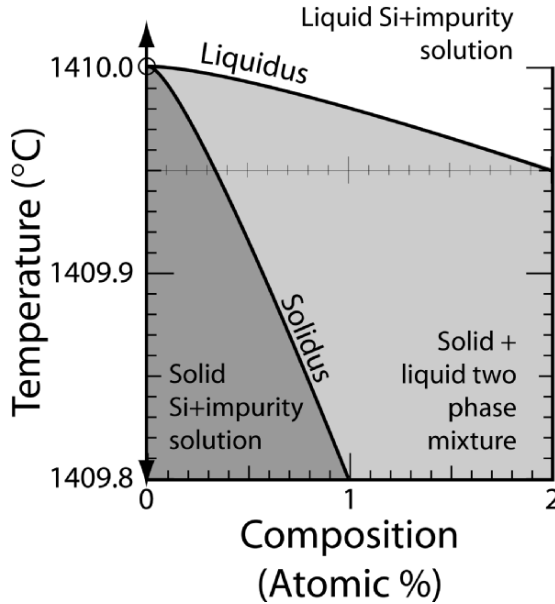
Suppose that you are growing a Si crystal by the Czochralski method from a liquid containing 2 at.% of the impurity with the corresponding phase diagram above.

   a) What temperature would you expect in the solid at the liquid interface in equilibrium? (You need not explain.)
   b) Calculate the segregation ratio for this impurity at this temperature.
   c) What is the effect on the temperature of the liquid near the interface of a chemically-altered liquid formed by rejection of the impurity into the liquid from the solidifying solid? (Explain in a sentence or two.)
   d) Briefly, how do you perform Czochralsky growth on a high vapor pressure material such as GaAs?

7. As an engineer at MegaJoule Industries your task is to grow bulk crystals by the Czochralsky process. One of the Czochralsky pullers has an air leak loud enough to hear during pump down. The worst part of the problem is that the leak cannot be fixed without replacing the entire unit. Because the whole growth facility is being replaced in a year it is not practical to fix this puller now. You find that the grown crystals contain both carbon and oxygen impurities as a result of the leak.

a) Could you remove the carbon that enters through the air leak by repeatedly melting and refreezing the boule? (This process is called zone refining and takes advantage of the tendency of some impurities to segregate from the solid to the liquid.) Explain briefly. (Segregation coefficient $C_s/C_l = 0.07$)

b) Would you expect the concentration of oxygen to increase or decrease in the Boule during this process if a small amount of oxygen was present in the zone refiner as residual gas? (Segregation coefficient $C_s/C_l = 1.25$)

c) In problem 6 above, there is a sketch of the phase diagram for Si and an impurity very close to the Si melting point and near pure Si. Make a similar schematic sketch for Si and O given the segregation coefficient in part (b) above.

4.9 SUGGESTED READINGS AND REFERENCES

Suggested Readings:

Cullity, B.D., *Elements of X-ray Diffraction*, 2nd edition. Reading: Addison-Wesley, 1978.

Shackelford, James F., *Introduction to Materials Science for Engineers*. New Jersey: Prentice Hall, 1996.

Shewmon, Paul G., *Diffusion in Solids* 2nd ed. Warrendale: TMS, 1989.

Shimura, Fumio, *Semiconductor Silicon Crystal Technology*. San Diego: Academic Press, 1989.

Van Vechten, J.A., "A Simple Man's View of the Thermochemistry of Semiconductors" in *Handbook on Semiconductors*, ed. T.S. Moss, Vol. 3, *Materials, Properties, and Preparation*, ed. S.P. Keller, North Holland, Amsterdam, 1980, Chapter 1.

Villars, P., Prince, A., and Okamoto, H., *Handbook of Ternary Alloy Phase Diagrams*, v. 7. Metals Park: ASM International, 1995.

References:

[1]   Villars, P., Prince, A., and Okamoto, H., *Handbook of Ternary Alloy Phase Diagrams*, v. 7. Metals Park: ASM International, 1995.

[2]   Gödecke, T., Haalboom, T., and Ernst, F., "Phase equilibria of Cu-In-Se. I. Stable states and nonequilibrium states of the In_2Se_3-Cu_2Se subsystem." *Zeitschrift Für Metallkunde,* 2000; 91: 622-34.