CHAPTER 1

Physics of Bulk Solids

In this text we will be discussing the physics and chemistry of nanostructures. The materials used to form these structures generally have bulk properties that become modified when their sizes are reduced to the nanorange; the present chapter presents background material on bulk properties of this type. Much of what is discussed here can be found in a standard text on solid-state physics.[1–4]

1.1. STRUCTURE

1.1.1. Size Dependence of Properties

Many properties of solids depend on the size range over which they are measured. Microscopic details become averaged when investigating bulk materials. At the macro- or large-scale range ordinarily studied in traditional fields of physics such as mechanics, electricity and magnetism, and optics, the sizes of the objects under study range from millimeters to kilometers. The properties that we associate with these materials are averaged properties, such as the density and elastic moduli in mechanics, the resistivity and magnetization in electricity and magnetism, and the dielectric constant in optics. When measurements are made in the micrometer or nanometer range many properties of materials change, such as mechanical, ferroelectric, and ferromagnetic properties. The aim of the present book is to examine characteristics of solids at the next lower level of size, namely, the nanoscale level, perhaps from 1 to 100 nm. Below this there is the atomic scale near 0.1 nm, followed by the nuclear scale near a femtometer (10^{-15} m). In order to understand properties at the nanoscale it is necessary to know something about the corresponding properties at the macroscopic and mesoscopic scales, and the present chapter aims to provide some of this background.

Many important nanostructures are composed of the group IV elements Si or Ge, types III–V semiconducting compounds such as GaAs or types II–VI semiconducting materials such as CdS, so these semiconductor materials will be used to illustrate some of the bulk properties that become modified when their dimensions are reduced

The Physics and Chemistry of Nanosolids. By Frank J. Owens and Charles P. Poole, Jr.

to the nanometer range. These Roman numerals IV, III, V, and, so on refer to columns of the periodic table. Tabulations of various bulk properties of these semiconductors are found in Appendix B.

1.1.2. Crystal Structures

Most solids are crystalline with their atoms arranged on sites of a regular lattice structure. They possess "long-range order" because this regularity extends throughout the entire crystal. In contrast to this amorphous materials such as glass and wax lack long-range order, but they have "short-range order," which means that the local environment of each atom is similar to that of other equivalent atoms, but this regularity does not persist over appreciable distances. In other words, each atom of a particular type might have, for example, six nearest neighbors, and these neighboring atoms might be at positions that approximate an octahedral configuration. Liquids also have short-range order, but lack long-range order, and this short-range order is undergoing continual rearrangements as a result of Brownian motion. Gases lack both long-range and short-range order. Their constituent molecules undergo rapid translational motion in all directions, so the disorder is continually rearranging at a rapid rate.

Figure 1.1 shows the five regular arrangements of lattice points that can occur in two dimensions, namely, the square (a), primitive rectangular (b), centered rectangular (c), hexagonal (d), and oblique (e) kinds. These arrangements are called *Bravais lattices*. The general or oblique Bravais lattice has two unequal lattice constants $a \neq b$ and an arbitrary angle θ between them. For the perpendicular case when $\theta = 90^\circ$, the lattice becomes rectangular, and if in addition $a = b$, the lattice is called *square*. For the special case $a = b$ and $\theta = 60^\circ$, the lattice is hexagonal, formed from equilateral triangles. Each lattice has a unit cell, indicated in the figures, which can replicate throughout the plane and generate the lattice.

A crystal structure is formed by associating with a lattice a regular arrangement of atoms or molecules. Figure 1.2 presents a two-dimensional crystal structure based on a primitive rectangular lattice containing two diatomic molecules A–B in each unit cell. A single unit cell can generate the overall lattice.

In three dimensions there are three lattice constants, a, b, and c, and three angles: α between b and c, β between a and c, and γ between lattice constants a and b. There are 14 Bravais lattices, ranging from the lowest symmetry triclinic type in which all three lattice constants and all three angles differ from each other ($a \neq b \neq c$

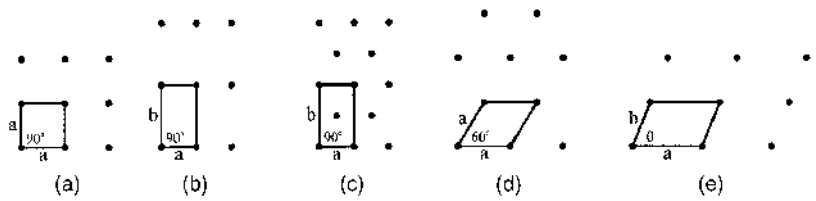

Figure 1.1. The five Bravais lattices that occur in two dimensions, with the unit cells indicated: (a) square; (b) primitive rectangular; (c) centered rectangular; (d) hexagonal; (e) oblique.

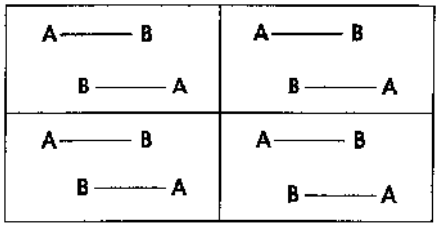

Figure 1.2. Sketch of a two-dimensional crystal structure based on a primitive rectangular lattice containing two diatomic molecules A–B in each unit cell.

and $\alpha \neq \beta \neq \gamma$), to the highest-symmetry cubic case in which all the lattice constants are equal and all the angles are 90° ($a = b = c$ and $\alpha = \beta = \gamma = 90°$). There are three Bravais lattices in the cubic system: a primitive or simple cubic (sc) lattice in which the atoms occupy the eight apices of the cubic unit cell, as shown in Fig. 1.3a; a body-centered cubic (bcc) lattice with lattice points occupied at the apices and in the center of the unit cell, as indicated in Fig. 1.3b; and a face-centered cubic (fcc) Bravais lattice with atoms at the apices and in the centers of the faces, as shown in Fig. 1.3c.

In two dimensions the most efficient way to pack identical circles (or spheres) is the equilateral triangle arrangement shown in Fig. 1.4a, corresponding to the hexagonal Bravais lattice of Fig. 1.1d. A second hexagonal layer of spheres can be positioned on top of the first to form the most efficient packing of two layers, illustrated in Fig. 1.4b. For efficient packing, the third layer can be placed either above the first layer with an atom at the location indicated by *T* or in the third possible arrangement with an atom above the position marked by X on the figure. In the first case a hexagonal lattice structure called *hexagonal close-packed* (hcp) is generated, and in the second case a fcc lattice results. The former is easy to identify in a unit cell, but the latter is not so easy to visualize while looking at a unit cell since the close-packed planes are oriented perpendicular to the [111] direction.

In three dimensions the unit cell of the fcc structure is the cube shown in Fig. 1.3c, which has a side (i.e., lattice constant) a and volume a^3. It has six face-centered atoms, each shared by two unit cells, and eight apical atoms, each shared by eight unit cells, corresponding to a total of four for this individual unit cell. Nearest

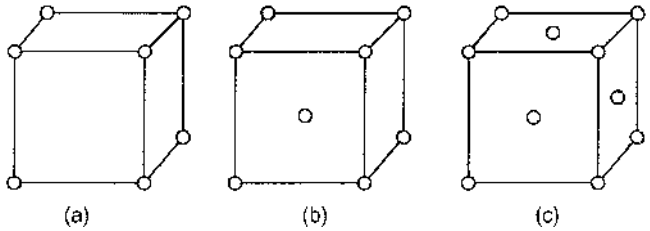

Figure 1.3. Unit cells of the three cubic Bravais lattices: (a) simple cubic (sc); (b) body-centered cubic (bcc); (c) and face-centered cubic (fcc).

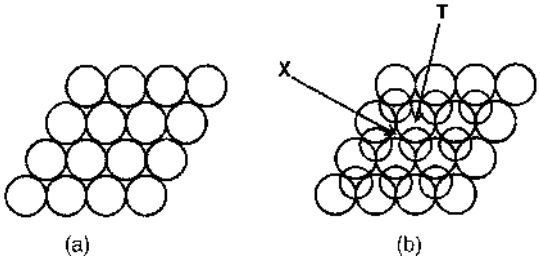

Figure 1.4. Close packing of spheres on a flat surface (a) for a monolayer and (b) with a second layer added. The circles of the second layer are drawn smaller for clarity. The location of an octahedral site is indicated by X, and the position of a tetrahedral site is designated by T in (b).

neighbors are a distance $a/(2)^{1/2}$ apart, so the atomic radius is defined as one-half of this value, namely, $r_A = a/2(2)^{1/2}$. Since there are four atoms per unit cell the volume density is $\rho_V = 4/a^3$. Each spherical atom has the individual volume $4\pi\, r_A^3/3$, so the packing fraction or percentage of the unit cell volume occupied by the atoms is $\pi/3(2)^{1/2} = 0.7406$. Figure 1.5 illustrates an alternate way to present the unit cell of the fcc lattice. There are one atom in the center, and there are 12 atoms at the edges. Each edge atom is shared by four unit cells so it counts as one-fourth of this particular cell, and the centrally located atom is not shared, corresponding to the expected total of four for this particular unit cell.

If only the 13 atoms shown in Fig. 1.5 are present, then the configuration constitutes the nanoparticle displayed in Fig. 1.6; this is discussed in the next section. This nanoparticle may be assumed to consist of three layers of two-dimensional close-packed planes perpendicular to its body diagonal or [111] axis, as indicated in Fig. 1.7.

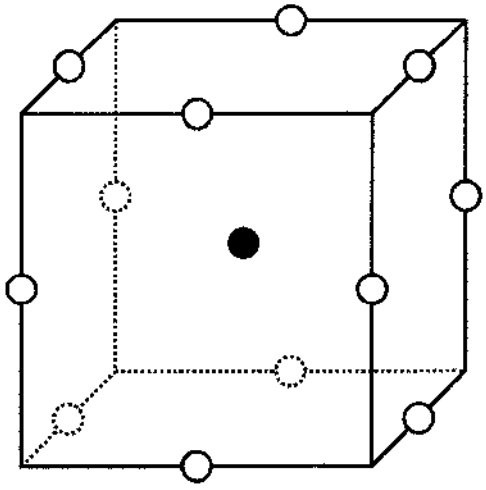

Figure 1.5. Face-centered cubic unit cell showing the 12 nearest-neighbor atoms that surround the atom (darkened circle) in the center.

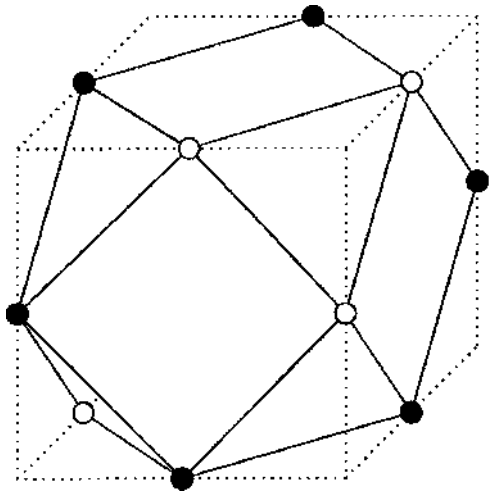

Figure 1.6. A 13-atom nanoparticle set in its fcc unit cell, showing the shape of the 14-sided polyhedron associated with the nanoparticle. The three open circles at the upper right correspond to the top layer, the six solid circles plus the atom (not pictured) in the center of the cube constitute the middle hexagonal layer, and the open circle at the lower left corner of the cube is one of the three atoms at the bottom layer of the cluster.

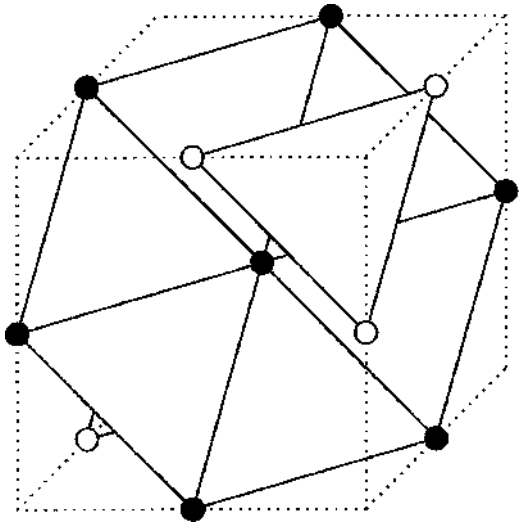

Figure 1.7. A 13-atom nanocluster set in its fcc centered cubic unit cell, delineating the planar configurations of the 13 included atoms. The 7-atom (darkened circles) plane in the center lies between the 3-atom (unshaded circles) plane at the upper right and the 3-atom partly hidden plane at the rear. Figure 1.6 provides a similar perspective.

There is a centrally located seven-atom close-packed plane (blackened circles), sandwiched between a pair of three-atom close-packed planes (unshaded circles). Two of the three atoms of the lower such plane are obscured from view on the figure. The distance between adjacent planes is one-third of the length $(3)^{1/2}$ a of the body diagonal, or $a/(3)^{1/2}$. The apex of the cube along this body diagonal is the distance $a/2(3)^{1/2}$ above the upper close-packed (three-atom) plane, and it is the position of an octahedral site equidistant between two close-packed planes. The atom configurations presented in Fig. 1.7 provide an easy way to visualize the relationships between fcc close-packed planes.

In the three-dimensional arrangements of close-packed spheres there are spaces or sites between the spheres where smaller atoms can reside. These sites exist between pairs of close-packed layers, so they are independent of the manner in which a third layer is added. In other words, they are the same for fcc and hcp lattices.

The point marked by X on Fig. 1.4b is called an *octahedral site* since it is equidistant from the three spheres O below it, and from the three spheres O above it, and these six nearest neighbors constitute the apices of an imaginary octahedron. An atom A at this site has the local coordination AO_6. The radius a_{oct} of this octahedral site is

$$a_{oct} = \frac{1}{4}(2 - (2)^{1/2})a = ((2)^{1/2} - 1)a_0 = 0.41411a_0 \tag{1.1}$$

where a is the lattice constant and a_0 is the sphere radius. The number of octahedral sites is equal to the number of spheres. There are also smaller sites called *tetrahedral sites*, labeled T on the figure that are equally distant from four nearest-neighbor spheres, one below and three above, corresponding to AO_4 for the local coordination. These four spheres define an imaginary tetrahedron surrounding an atom A located at the site. This is a smaller site since its radius a_T is

$$a_T = \frac{1}{4}((3)^{1/2} - (2)^{1/2})a = [(3/2)^{1/2} - 1]a_0 = 0.2247a_0 \tag{1.2}$$

There are twice as many tetrahedral sites as spheres in the structure. Many diatomic oxides and sulfides such as MgO, MgS, MnO, and MnS possess larger oxygen or sulfur anions in a perfect fcc arrangement with the smaller metal cations located at octahedral sites. This is an NaCl lattice, where we use the term *anion* for a negative ion (e.g., Cl^-), and cation for a positive ion (e.g., Na^+). The mineral spinel $MgAl_2O_4$ has a face-centered arrangement of divalent oxygens O^{2-} (radius 0.132 nm) with the Al^{3+} ions (radius 0.051 nm) occupying one-half of the octahedral sites and Mg^{2+} (radius 0.066 nm) located in one-eighth of the tetrahedral sites in a regular manner. In some spinels the oxygens deviate somewhat from a perfect fcc arrangement by moving slightly toward or away from the tetrahedral sites.

1.1.3. Face-Centered Cubic Nanoparticles

Most metals in the solid-state form close-packed lattices; thus Ag, Al, Au, Co, Cu, Pb, Pt, and Rh, as well as the rare gases Ne, Ar, Kr, and Xe, are fcc, and Mg, Nd, Os, Re, Ru, Y, and Zn are hcp. Some other metallic atoms crystallize in the more loosely packed bcc lattice, and a few such as Cr, Li, and Sr crystallize in all three structure types, depending on the temperature. An atom in each of the two close-packed lattices has 12 nearest neighbors. Figure 1.5 shows the 12 neighbors that surround an atom (darkened circle) located in the center of a cube for a fcc lattice. In Fig. 1.7 we show a redrawn version of the fcc unit cell of Fig. 1.6 that clarifies the layered structure of the atoms. The topmost layer is an equilateral triangle of atoms (open circles); the next layer below is a regular hexagon of atoms (blackened circles); and the bottom layer, another equilateral triangle of open circle atoms, is mostly hidden from view. These 13 atoms constitute the smallest theoretical nanoparticle for an fcc lattice. More precisely it should probably be called a *nanocluster*. Figure 1.6 shows the 14-sided polyhedron, called a *dekatessarahedron*, that is generated by connecting the nearest-neighbor atoms. This polyhedron has six square faces and eight equilateral triangle faces. Figure 1.8 shows this polyhedron viewed from the [111] or top direction with the coordinate axis perpendicular to each side indicted. The square sides are perpendicular to the x axis [100], the y axis [010], and the z axis [001], respectively, and the triangular faces are perpendicular to the indicated body diagonal directions. Figure 1.7 clarifies the layering scheme by connecting atoms only within two-dimensional close-packed layers. The three open circles in the triangular layer of Fig. 1.7 are the three atoms in the top layer of Fig. 1.8. Likewise, the seven darkened

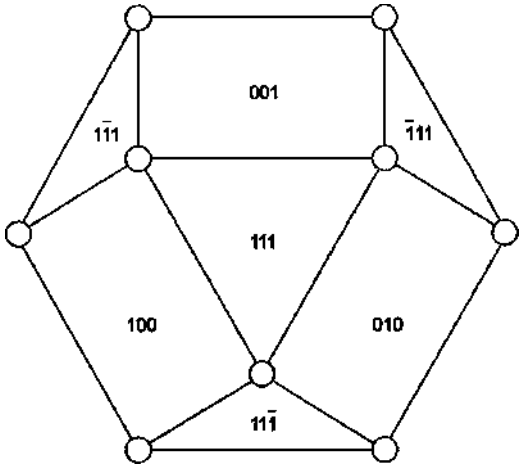

Figure 1.8. Perspective along the [111] direction of the faces or facets of the 13-atom fcc nanoparticle of Fig. 1.6. The three visible square faces are perpendicular to the x direction [100], the y direction [010], and the z direction [001], respectively. The four equilateral triangle faces are perpendicular to body diagonal directions, as indicated.

circles, including the atom in the center of the cube of Fig. 1.7, constitute the periphery of Fig. 1.8.

To construct the next-larger 55-atom fcc nanoparticle, a layer of 30 atoms is added around the three layers of Fig. 1.7, and six-atom triangular arrays are added to the top and bottom, as explained in Appendix C. The process is repeated to form the higher-level nanoparticles with the magic numbers of atoms listed in Table 1.1. A perspective sketch of the atoms on the surfaces of the 147-atom nanoparticle viewed along the [111] axis is provided in Fig. 1.9, with the crystallographic planes labeled by their indices. There are six square faces with square planar arrays of 16 atoms each, and eight equilateral triangle faces with planar hcp arrays of 10 atoms each. The atoms on all of these triangular and square faces have the nearest-neighbor separation $a/2^{1/2}$, where a is the lattice constant. Larger fcc nanoparticles with magic numbers 309, 561, ... of atoms have analogous arrays of atoms on their faces, with the numbers of surface atoms involved in each case listed in Table C.1. The numbers of atoms in the various planes of these nanoparticles are tabulated in Table C.2. The nanoparticles pictured in Fig. 3.18 (Fig. 4.15 of first edition), and in Fig. 2.13 (Fig. 3.13 of the first edition) clearly display both the planar square and the planar hexagonal arrays of atoms corresponding to those on the faces of Fig. 1.9.

TABLE 1.1. Diameter, Total Number of Atoms, Number on Surface, and Percentage on Surface of Rare Gas or Metallic Nanoparticles with FCC Close-Packed Structures and a Structural Magic Number of Atoms[a]

Shell	Diameter	Total Number	Number on Surface	Percentage on Surface
1	d	1	1	100
2	$3d$	13	12	92.3
3	$5d$	55	42	76.4
4	$7d$	147	92	62.6
5	$9d$	309	162	52.4
6	$11d$	561	252	44.9
7	$13d$	923	362	39.2
8	$13d$	1415	492	34.8
9	$17d$	2057	642	31.2
10	$19d$	2869	812	28.3
11	$21d$	3871	1002	25.9
12	$23d$	5083	1212	23.8
25	$49d$	4.90×10^4	5.76×10^3	11.7
50	$99d$	4.04×10^5	2.40×10^4	5.9
75	$149d$	1.38×10^6	5.48×10^4	4.0
100	$199d$	3.28×10^6	9.80×10^4	3.0

[a]The diameters d in nanometers for some representative fcc atoms are Ag 0.289, Al 0.286, Ar 0.372, Au 0.288, Cu 0.256, Fe 0.254, Kr 0.404, Pb 0.350, and Pd 0.275.

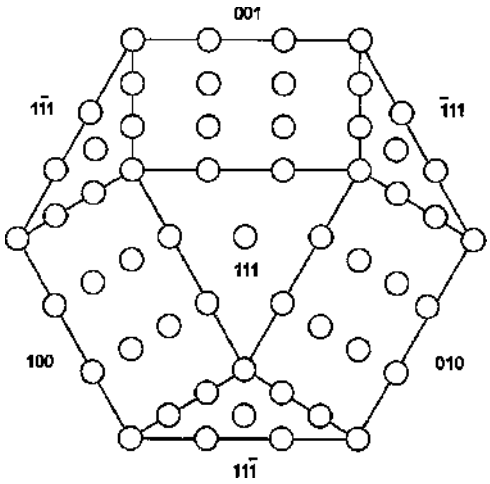

Figure 1.9. Arrays of atoms on triangular and square faces of the 147-atom fcc nanoparticle. The directions perpendicular to each face are indicated.

1.1.4. Large Face-Centered Cubic Nanoparticles

Larger fcc nanoparticles with the same polyhedral shape are obtained by adding more planes or layers, and the sequence of numbers in the resulting particles, $N =$ 1,13,55,147,309, . . . , listed in Table 1.1, are called *structural magic numbers*. For n layers the number of atoms N in this fcc nanoparticle is given by the formula

$$N = (1/3)[10n^3 - 15n^2 + 11n - 3] \tag{1.3}$$

and the number of atoms on the surface N_{surf} is

$$N_{\text{surf}} = 10n^2 - 20n + 12 \tag{1.4}$$

Equation (1.3) is valid for all n, but Eq. (1.4) is valid only for $n > 1$. For each value of n Table 1.1 lists the number of atoms on the surface, as well as the percentage of atoms on the surface. The table also lists the diameter of the each nanoparticle, which is given by the expression $(2n - 1)d$, where d is the distance between the centers of nearest-neighbor atoms and $d = a/(2),^{1/2}$, where a is the lattice constant.

Purely metallic fcc nanoparticles such as Au_{55} tend to be very reactive and have short lifetimes. They can be ligand-stabilized by adding atomic groups between their atoms and on their surfaces. The Au_{55} nanoparticle has been studied in the ligand-stabilized form $Au_{55}(PPh_3)_{12}Cl_6$, which has the diameter of $\approx$1.4 nm, where PPh_3 is an organic group. Further examples are the magic number nanoparticles Pt_{309}(1,10-phenantroline)$_{36}O_{30}$, and Pd_{561}(1,10-phenantroline)$_{36}O_{200}$.

The magic numbers that we have been discussing are called structural magic numbers because they arise from minimum-volume, maximum-density nanoparticles

that approximate a spherical shape, and have close-packed structures characteristic of a bulk solid. These magic numbers take no account of the electronic structure of the constituent atoms in the nanoparticle. Sometimes the predominant factor in determining the minimum-energy structure of small nanoparticles is the interactions between the valence electrons of the constituent atoms with an averaged molecular potential, so that the electrons occupy orbital levels associated with this potential. Atomic cluster configurations in which these electrons fill closed shells are especially stable, and constitute electronic magic numbers. Their atomic structures differ from the fcc arrangement, as will be discussed in Section 3.2.

When mass spectra were recorded for sodium nanoparticles Na_N it was found that mass peaks corresponding to the first 15 electronic magic numbers $N = 3,9,20,36,61,\ldots$ were observed for cluster sizes up to $N = 1220$ atoms ($n = 15$), and fcc structural magic numbers starting with $N = 1415$ for $n = 8$ were observed for larger sizes.[5,6] The mass spectral data are plotted versus the cube root of the number of atoms $N^{1/3}$ on Fig. 1.10, and it is clear that the lines from both sets of magic numbers are approximately equally spaced, with the spacing between the structural magic numbers about 2.6 times that between the electronic ones. This result is evidence that small clusters tend to satisfy electronic criteria, and large structures tend to be structurally determined.

1.1.5. Tetrahedrally Bonded Semiconductor Structures

Types III–V and II–VI binary semiconducting compounds, such as GaAs and ZnS, respectively, crystallize with one atom situated on a fcc sublattice at positions 000, $\frac{1}{2}\frac{1}{2}0$, $\frac{1}{2}0\frac{1}{2}$, and $0\frac{1}{2}\frac{1}{2}$, and the other atom on a second fcc sublattice displaced from the first by the amount $\frac{1}{4}\frac{1}{4}\frac{1}{4}$ along the body diagonal, as shown in Fig. 1.11b. This is called the *zinc blende* or ZnS structure. It is clear from the figure that each Zn atom (white sphere) is centered in a tetrahedron of S atoms (black spheres), and likewise each S has four similarly situated Zn nearest neighbors. The small half-sized, dashed-line cube delineates one such tetrahedron. The same structure would result if the Zn and S atoms were interchanged. Figure 1.12 presents a more realistic sketch of the unit cell of ZnS with the small zinc atoms centered in the tetrahedral sites of

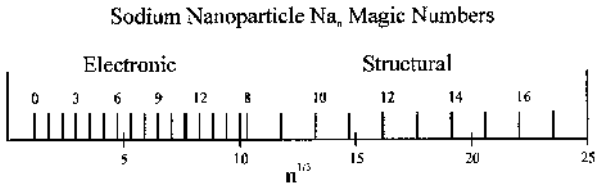

Figure 1.10. Dependence of the observed mass spectra lines from Na_N nanoparticles on the cube root $N^{1/3}$ of the number of atoms N in the cluster. The lines are labeled with the index n of their respective electronic and structural magic numbers. [Adapted from T. P. Martin, T. Bergmann, H. Gohlich, and T. Lange, *Chem. Phys. Lett.* **172**, 209 (1990).]

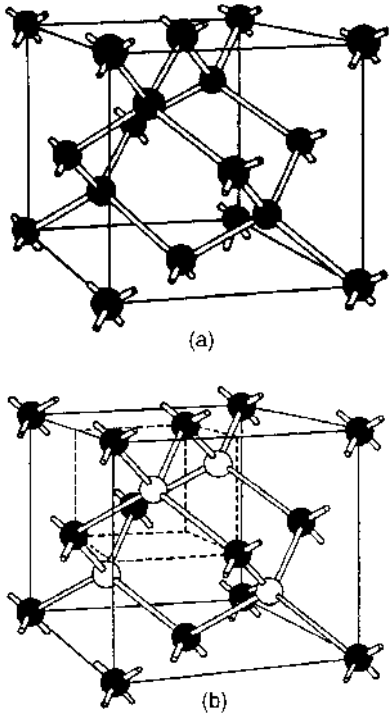

Figure 1.11. Unit cell of the diamond structure (a) containing only one type of atom, and corresponding unit cell of the zinc blende structure (b) containing two atom types. The rods represent the tetrahedral bonds between nearest-neighbor atoms. The small dashed-line cube in (b) delineates a tetrahedron. (From G. Burns, *Solid State Physics*, Academic Press, Boston, 1985, p. 148.)

the larger sulfur atoms. In contrast to this arrangement, the atoms in the semiconductor gallium arsenide have comparable radii $r_{Ga} = 0.122$ nm and $r_{As} = 0.124$ nm, and the lattice constant $a = 0.565$ nm is close to the value $4(r_{Ga} + r_{As})/(3)^{1/2} = 0.568$ nm, as expected, where data were used from Tables B.1 and B.2 (of Appendix B).

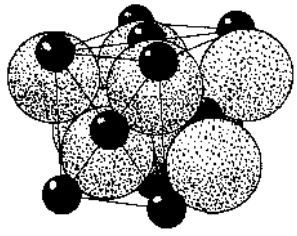

Figure 1.12. Packing of larger sulfur atoms and smaller zinc atoms in the zinc blende (ZnS) structure. Each atom is centered in a tetrahedron of the other atom type. (From R. W. G. Wyckoff, *Crystal Structures*, Vol. 1, Wiley, New York, 1963, p. 109.)

The elements Si and Ge crystallize in this same structure, in which the Si (or Ge) atoms occupy all the sites on the two sublattices, so there are eight identical atoms in the unit cell. This atom arrangement, sketched in Fig. 1.11a, is called the "diamond structure." Both Si and Ge have a valence of 4, so from bonding considerations it is appropriate for each to be bound to four other atoms in the shape of a regular tetrahedron.

In Table B.1 we see the lattice constants a listed for various compounds with the zinc blende structure, and Table B.2 provides the crystal radii of their monatomic lattices in which the atoms are uncharged, as well as the ionic radii for ionic compounds in which the atoms are charged. We see from Table B.2 that the negative anions are considerably larger than the positive cations, in accordance with the sketch of the unit cell presented in Fig. 1.12, and this size differential is greater for the III–V compounds than for II–VI compounds. However, these size changes for the negative and positive ions tend to balance each other so that the III–V compounds have the same range of lattice constants as do the II–VI compounds, with Si and Ge also in this range. Table B.4 gives the molecular masses, and Table B.5 gives the densities of these semiconductors. The three tables, Tables B.1, B.4, and B.5, show a regular progression in the values as one goes from left to right in a particular row, and as one goes from top to bottom in a particular column. This occurs because of the systematic increase in the size of the atoms in each group with increasing atomic number, as indicted in Table B.2.

There are two simple models for representing these AC binary compound structures. For an ionic model the lattice $A^{n-}C^{n+}$ consists of a fcc arrangement of the large anions A^{n-} with the small cations C^{n+} located in the tetrahedral sites of the anion fcc lattice. If the anions touch each other, their radii has the value $r = a/2(2)^{1/2}$, where a is the lattice parameter, and the radius a_T of the tetrahedral site $r_{tetr} = 0.2247\ a$ is given by Eq. (1.2). This is the case for the very small Al^{3+} cation in the AlSb structure. In all other cases the cations in Table B.2 are too large to fit in the tetrahedral site, so they push the larger anions further apart, and the latter no longer touch each other, in accordance with Fig. 1.12. In a covalent model for the structure consisting of neutral atoms A and C the atom sizes are comparable, as the data in Table B.2 indicate, and the structure resembles that of Si or Ge. Comparing these two models, we note that the distance between atom A at lattice position 000 and its nearest neighbor C at position $\frac{1}{4}\frac{1}{4}\frac{1}{4}$ is equal to $\frac{1}{4}(3)^{1/2}a$, and in Table B.3 we compare this crystallographically evaluated distance with the sums of radii of ions A^{n-}, C^{n+} from the ionic model, and with the sums of radii of neutral atoms A and C of the covalent model using the data of Table B.2. We see from the results in Table B.3 that neither model fits the data in all cases, but the neutral atom covalent model is closer to agreement, especially for ZnS, GaAs, and CdS. For comparison purposes we also list corresponding data for several alkali halides NaCl, KBr, and RbI, and alkaline-earth chalcogenides CaS and SrSe, which also crystallize in the cubic rock salt or NaCl structure. We see that all of these compounds fit the ionic model very well. In these compounds each atom type forms a fcc lattice, with the atoms of one fcc lattice located at octahedral sites of the other lattice. The octahedral site has the radius $r_{oct} = 0.41411\ a_0$ given by Eq. (1.1), which is larger than the tetrahedral one, $r_{tetr} = 0.2247\ a$, of Eq. (1.2).

Since the alkali halide and alkaline-earth chalcogenide compounds fit the ionic model so well, it is significant that neither model fits the structures of the semiconductor compounds. The extent to which the semiconductor crystals exhibit ionic or covalent bonding is not clear from crystallographic data. If the wavefunction describing the bonding is written in the form

$$\Psi = a_{\text{cov}}\psi_{\text{cov}} + a_{\text{ion}}\psi_{\text{ion}} \tag{1.5}$$

where the coefficients of the covalent and ionic wavefunction components are normalized

$$a_{\text{cov}}{}^2 + a_{\text{ion}}{}^2 = 1 \tag{1.6}$$

then $a_{\text{cov}}{}^2$ is the fractional covalency and $a_{\text{ion}}{}^2$ is the fractional ionicity of the bond. A chapter in a recent book by Karl Boer[7] tabulates the effective charges e^* associated with various II–VI and III–V semiconducting compounds, and this effective charge is related to the fractional covalency by the expression

$$a_{\text{cov}}{}^2 = \frac{8 - N + e^*}{8} \tag{1.7}$$

where $N = 2$ for II–VI and $N = 3$ for III–V compounds. The fractional charges all lie in the range from 0.43 to 0.49 for the compounds under consideration. Using the e^* tabulations in the Boer book and Eq. (1.7), we obtain the fractional covalencies of $a_{\text{cov}}{}^2 \sim 0.81$ for all the II–VI compounds, and $a_{\text{cov}}{}^2 \sim 0.68$ for all the III–V compounds listed in Tables B.1, B.4, B.5, and so on. These values are consistent with the better fit of the covalent model to the crystallographic data for these compounds.

We conclude this section with some observations that will be of use in later chapters. Table B.1 shows that the typical compound GaAs has the lattice constant $a = 0.565$ nm, so the volume of its unit cell is $0.180\,\text{nm}^3$, corresponding to about 22 of each atom type per cubic nanometer. The distances between atomic layers in the [100], [110], and [111] directions are, respectively, $a/2 = 0.28$ nm, $a/(2)^{1/2} = 0.40$ nm, and $a/(3)^{1/2} = 0.33$ nm, for GaAs. The various III–V semiconducting compounds under discussion form mixed crystals over broad concentration ranges, as do the group of II–VI compounds. In a mixed crystal of the type $\text{In}_x\text{Ga}_{1-x}\text{As}$, it is ordinarily safe to assume that Vegard's law is valid, whereby the lattice constant a scales linearly with the concentration parameter x. As a result, we have the following relationships

$$\begin{aligned} a(x) &= a(\text{GaAs}) + [a(\text{InAs}) - a(\text{GaAs})]x \\ &= 0.565 + 0.039x \end{aligned} \tag{1.8}$$

where $0 < x < 1$. In the corresponding expression for the mixed semiconductor $\text{Al}_x\text{Ga}_{1-x}\text{As}$ the term $\leq 0.003x$ replaces the term $+0.039x$, so the fraction of lattice

mismatch $2|a_{\mathrm{AlAs}} - a_{\mathrm{GaAs}}|/(a_{\mathrm{AlAs}} + \mathrm{a}_{\mathrm{GaAs}}) = 0.0054 = 0.54\%$ for this system is quite minimal compared to that $2|a_{\mathrm{InAs}} - a_{\mathrm{GaAs}}|/(a_{\mathrm{InAs}} + a_{\mathrm{GaAs}}) = 0.066 = 6.6\%$ of the $\mathrm{In}_x\mathrm{Ga}_{1-x}\mathrm{As}$ system, as calculated from Eq. (1.8). Table B.1 of Appendix B gives the lattice constants a for various III–V and II–VI semiconductors with the zinc blende structure.

1.1.6. Lattice Vibrations

We have discussed atoms in a crystal as residing at particular lattice sites, but in reality they undergo continuous fluctuations in the neighborhood of their regular positions in the lattice. These fluctuations arise from the heat or thermal energy in the lattice, and become more pronounced at higher temperatures. Since the atoms are bound together by chemical bonds, the movement of one atom about its site causes the neighboring atoms to respond to this motion. The chemical bonds act like springs that stretch and compress repeatedly during the oscillatory motion. The result is that many atoms vibrate in unison, and this collective motion spreads throughout the crystal. Every type of lattice has its own characteristic modes or frequencies of vibration called *normal modes*, and the overall collective vibrational motion of the lattice is a combination or superposition of many, many normal modes. For a diatomic lattice like GaAs, there are low-frequency modes called *acoustic modes*, in which the heavy and light atoms tend to vibrate in phase or in unison with each other, and high-frequency modes called *optical modes*, in which they tend to vibrate out of phase.

A simple model for analyzing these vibratory modes is a linear chain of alternating atoms with a large mass M and a small mass m joined to each other by springs ($\sim$) as follows:

$$\sim m \sim M \sim m \sim M \sim m \sim M \sim m \sim M \sim$$

When one of the springs stretches or compresses by an amount Δx, a force is exerted on the adjacent masses with the magnitude $C\,\Delta x$, where C is the spring constant. As the various springs stretch and compress in step with each other, longitudinal modes of vibration take place in which the motion of each atom is along the string direction. Each such normal mode has a particular frequency ω and a wavevector $k = 2\pi/\lambda$, where λ is the wavelength, and the energy E associated with the mode is given by $E = \hbar\omega$. There are also transverse normal modes in which the atoms vibrate back and forth in directions perpendicular to the line of atoms. Figure 1.13 shows the dependence of ω on k for the low-frequency acoustic and the high-frequency optical longitudinal modes. We see that the acoustic branch continually increases in frequency ω with increasing wavenumber k, and the optical branch continuously decreases in frequency. The two branches have respective limiting frequencies given by $(2C/M)^{1/2}$ and $(2C/m)^{1/2}$, with an energy gap between them at the edge of the Brillouin zone $k_{\max} = \pi/a$, where a is the distance between atoms m and M

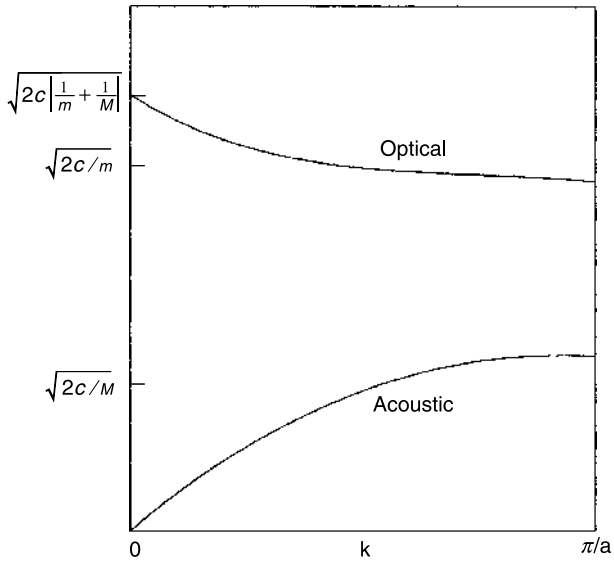

Figure 1.13. Dependence of the longitudinal normal-mode vibrational frequency ω on the wavenumber $k = 2\pi/\lambda$ for a linear diatomic chain of atoms with alternating masses $m < M$ having an equilibrium spacing a, and connected by bonds with spring constant C. [From C. P. Poole, Jr., *The Physics Handbook*, Wiley, New York, 1998, p. 53.)

at equilibrium. The Brillouin zone is a unit cell in wavenumber or reciprocal space, as will be explained in Section 1.3.2 later in this chapter. The optical branch vibrational frequencies are in the infrared region of the spectrum, generally with frequencies in the range from 10^{12} to 3×10^{14} Hz, and the acoustic branch frequencies are much lower. In three dimensions the situation is more complicated, and there are longitudinal acoustic (LA), transverse acoustic (TA), longitudinal optical (LO), and transverse optical (TO) modes.

The atoms in molecules also undergo vibratory motion, and a molecule containing N atoms has $3N - 6$ normal modes of vibration. Particular molecular groups such as hydroxyl —OH, amino —NH_2, and nitro —NO_2 have characteristic normal modes that can be used to detect their presence in molecules and solids.

The atomic vibrations that we have been discussing correspond to standing waves. This vibrational motion can also produce traveling waves in which localized regions of vibratory atomic motion travel through the lattice. Examples of such traveling waves are sound moving through the air, or seismic waves that start at the epicenter of an earthquake and travel thousands of miles to reach a seismograph detector that records the earthquake event many minutes later. Localized traveling waves of atomic vibrations in solids, called *phonons*, are quantized with the energy $\hbar\omega = h\nu$, where $\nu = \omega/2\pi$ is the frequency of vibration of the wave. Phonons play an important role in the physics of the solid state.

1.2. SURFACES OF CRYSTALS

1.2.1. Surface Characteristics

When we consider the bulk properties of macroscopic crystals, the surface layers play a negligible role. If a simple cubic crystal in the shape of a cube contains 10^{24} atoms, then each face of the cube will contain 10^{16} atoms, for a total of 6×10^{16} on the surface. This is less than one atom in 10^7 on the surface, so the bulk is not appreciably affected by the presence of the surface. On the other hand, when we consider particles in the nanoscale range of dimensions, the percentage of the atoms on the surface can be large. We see from Table 1.1 that an aluminum nanoparticle with a diameter $49d = 14.0$ nm contains 4.9×10^4 atoms, 11.7% of which are on the surface. The largest aluminum nanoparticle listed in the table has a diameter of 56.9 nm and contains 3.28 million atoms, 3% of which are on the surface. As a result, the surface can influence the bulk properties. The same is true of thin films with nanometer thicknesses. For example, consider a fcc thin film consisting of 100 close-packed layers (the [111] axis is perpendicular to the surface). Such a film made of copper would be 21 nm thick. Since there is both a top surface and a bottom surface, it follows that 2% of the atoms are in a surface layer.

We have been discussing surface layers as if they are identical to the layers in the bulk, and this is seldom the case. Within the crystal the bonding orbitals of all the atoms are satisfied, and each type of atom is in the same environment as its counterparts. In contrast to this, the surface atoms have dangling bonds that are not satisfied, and this deficiency of chemical bonding can be characterized quantitatively by a surface energy per unit area. This surface energy results from the work that must be done to cleave the crystal and break the dangling surface bonds while forming the surface. Sometimes a surface atom reconstruction takes place to lower this surface energy. For example, the [111] surface of a silicon (or germanium) crystal consists of a hexagonal array of Si atoms, each of which, shown unshaded, has a single vertical dangling bond indicated by a dot in Fig. 1.14. Each surface atom on the figure is bonded to three Si atoms shown shaded in the layer below the surface. The energy is lowered by the surface reconstruction illustrated in Fig. 1.15, in which surface atoms move together and bond to each other in pairs to accommodate and satisfy the broken dangling bonds. Another phenomenon that sometimes takes place is surface structure relaxation in which the outer layer of atoms moves slightly toward or a short distance away from the layer below. Contraction takes place with most metal surfaces. The silicon [111] surface layer illustrated in Figs. 1.14 and 1.15 contracts by about 25%, and the three interlayer spacings further below compensate for this by expanding between 1% and 5%.

Other phenomena can occur at the surface. Sometimes there are steps on the surface beyond which a new layer has been added. Various types of defects can occur, such as added atoms in an irregular manner, or missing atoms at lattice sites. Surface atoms can bond together in ways that differ from those in the bulk, as took place in the reconstruction pictured in Fig. 1.15. Occasionally surface atoms will undergo a two-dimensional phase transition.

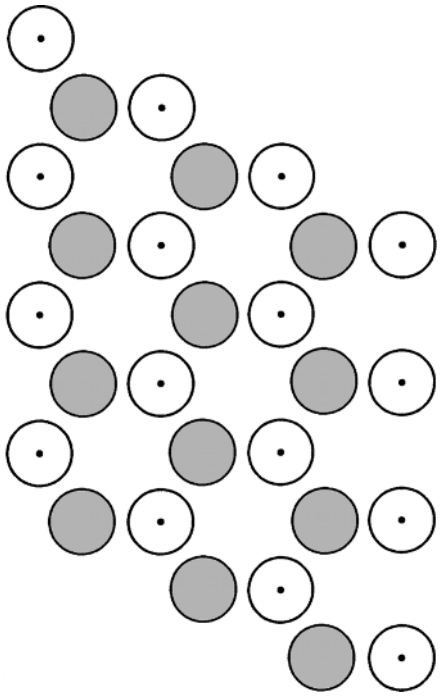

Figure 1.14. Sketch of the first two [111] outermost layers of the zinc blende structure showing unshaded sulfur atoms on the surface with vertical dangling bonds denoted by dots, and shaded zinc atoms one layer below the surface.

1.2.2. Surface Energy

Work must be done to cut a crystal into two pieces that expose a pair of new crystal faces or facets. The energy required to bring about this cleavage, called the *surface energy*, is proportional to the density of the various types of chemical bonds that are cut and transformed into dangling bonds. In the case of fcc crystals, each large atom has 12 large atom nearest neighbors to which it can be bonded. Each medium-sized atom in an octahedral site is bonded to six surrounding large atoms, and each small atom in a tetrahedral site is bonded to four large atoms. Which bonds break depends on the particular crystallographic plane where the cleavage occurs.

If the surface contains a density ρ_i of atoms of type i, each with n_i dangling bonds of energy $2\varepsilon_i$, then the surface energy density γ or energy per unit area is given by the following sum over the atom types

$$\gamma = \Sigma \rho_i n_i \varepsilon_i \tag{1.9}$$

where the factor of 2 takes into account the fact that the energy associated with a dangling bond is half of the corresponding chemical bond energy. If all of the atoms are

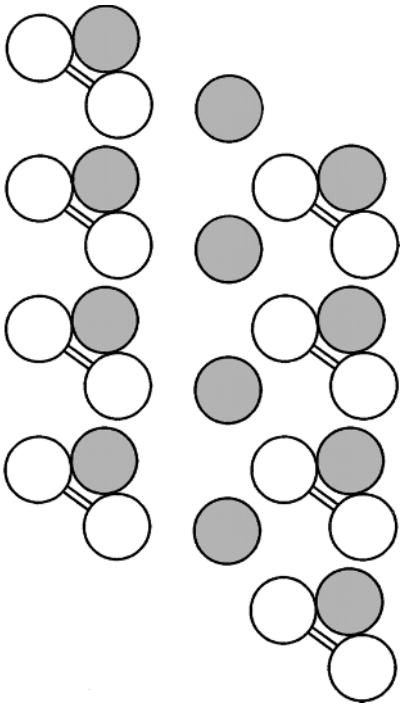

Figure 1.15. Sketch of the surface configuration of Fig. 1.14 after it has undergone a reconstructuring process during which surface sulfur atoms bond together in pairs.

identical as in the fcc and hcp lattices discussed above, this expression reduces to

$$\gamma = \rho n \varepsilon \tag{1.10}$$

In the next few sections we will evaluate this surface energy density for several special cases.

1.2.3. Face-Centered Cubic Surface Layers

We have been discussing how surface layers become modified relative to bulk layers in a crystal. Since many nanoparticles and nanofilms have a fcc structure, it will be instructive to consider the arrangement of the atoms on the surfaces of fcc cubic solids that are cleaved along particular crystallographic planes. Consider a solid with large atoms close-packed so that they touch all 12 nearest neighbors, and with the octahedral sites filled with medium-sized atoms in contact with their 6 nearest neighbors, and the tetrahedral sites occupied by small atoms in contact with their 4

nearest neighbors. The radii of these three types of atoms have the following values from Eqs. (1.1) and (1.2)

$$r_{\mathrm{L}} = 0.7072a \tag{1.11a}$$

$$r_{\mathrm{Oh}} = 0.4142a \tag{1.11b}$$

$$r_{\mathrm{T}} = 0.2247a \tag{1.11c}$$

where a is the lattice constant. Figure 1.16 shows an example of the octahedral sites occupied by smaller atoms, as happens in the NaCl structure. The zinc blende structure of Fig. 1.11 has half of the tetrahedral sites and none of the octahedral sites occupied. It is clear from these figures that some of the octahedral sites lie on the surface of the unit cell, while all of the tetrahedral ones are inside it. If the crystal is cleaved to expose the [100] face, it is evident from Fig. 1.16 that the surface will contain 50% larger and 50% smaller atoms, the latter occupying the octahedral sites of the former, as shown in Fig. 1.17 (i.e., 4 small half-atoms + 1 large atom + 4 large quarter-atoms). The [110] face displayed in Fig. 1.18 contains all three types of atoms. This can be seen by comparison with Figs. 1.11b and 1.16. The upper two tetrahedrally coordinated atoms on Fig. 1.18 are shaded to indicate that they are absent in the zinc blende structure. The [111] plane displayed in Fig. 1.4a contains only the main large atoms since none of the tetrahedral or octahedral sites are in this plane, as can be deduced from the positions of these sites specified in Fig. 1.4b.

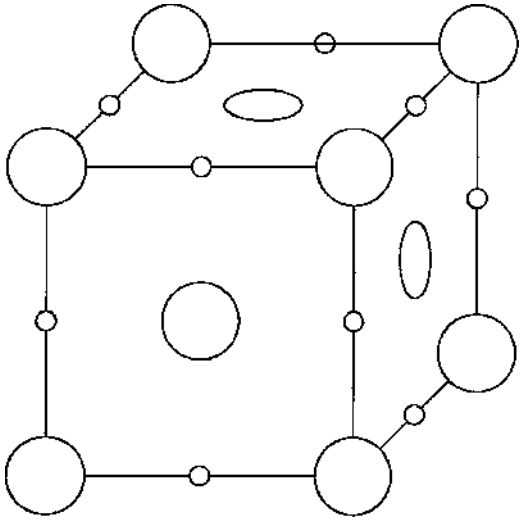

Figure 1.16. Unit cell consisting of fcc atoms (large circles) with the octahedral holes occupied by smaller atoms (small circles).

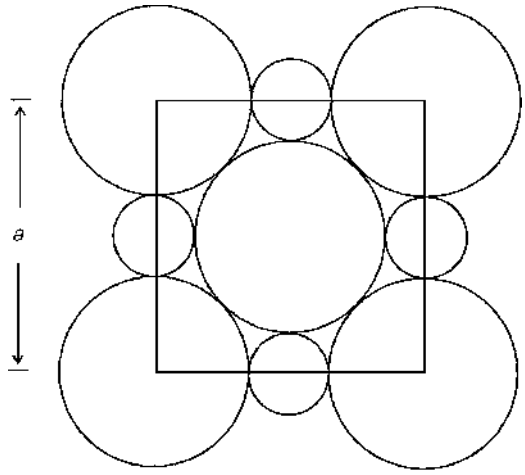

Figure 1.17. The [100] surface plane of the structure depicted in Fig. 1.16.

There are several special cases that can be clarified from the analysis above. If all of the atoms are of the same size and only octahedral sites are occupied, then the lattice reduces to a simple cubic one with the lattice constant $a/2$, and a unit cell one-eighth of the size. Since nanoparticles are rarely simple cubic, we will not discuss this case further. An important special case is when none of these extra sites are occupied so only the large atoms are present. This is the situation that

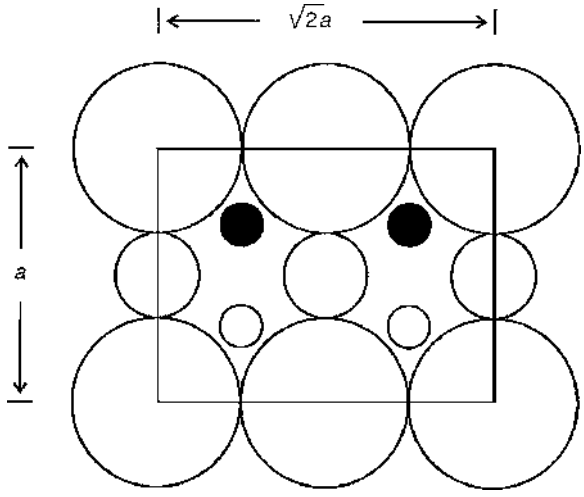

Figure 1.18. The [110] surface plane of a fcc structure formed by large atoms with the octahedral sites occupied by medium-sized atoms and the tetrahedral sites occupied by small atoms. The upper two small atoms shown shaded as well as the three medium-sized atoms are absent in the zinc blende structure.

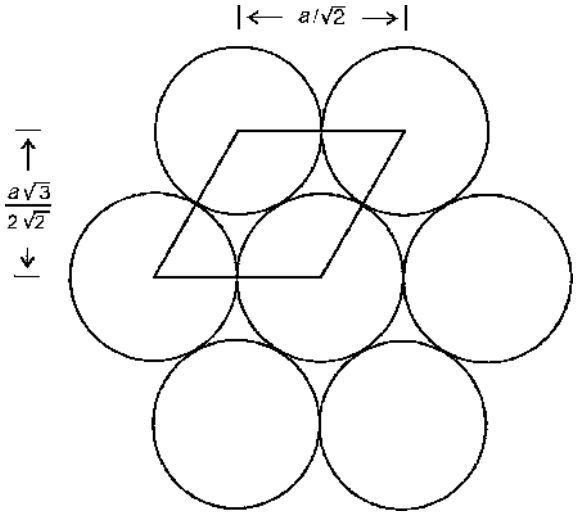

Figure 1.19. Surface unit cell of the close-packed lattice sketched in Fig. 1.4a.

corresponds to the fcc nanoparticles discussed in Section 1.1.3. The [100] face shown in Fig. 1.17 with the surface unit cell area a^2 contains two atoms, one entire atom in the center and four-quarters of an atom in the corners, so the atom density on the surface is $2/a^2$ (the smaller atoms in the figure are now absent). The center atom has four nearest neighbors inside the crystal face below it, four neighbors alongside it in the face, and four missing nearest neighbors above so four of its chemical bonds have been cut, and are now dangling bonds. This means that the number of dangling bonds per unit area is $8/a^2$, and the surface energy density from Eq. (1.10) is $\gamma = 8\varepsilon/a^2$. By similar reasoning, the [110] face unit cell shown in Fig. 1.18 with the surface unit cell area $(2)^{1/2}a^2$ contains two atoms, so the atom density on the surface is $(2)^{1/2}/a^2$. Each surface atom is bonded to five atoms below it and to two surface atoms, so it has five dangling bonds, for a dangling bond density of $5(2)^{1/2}/a^2$, and a surface energy density γ of $5(2)^{1/2}\varepsilon/a^2$. The [111] surface sketched in Fig. 1.19 is a little more complex to analyze. The rhombic unit cell outlined in the figure contains one atom and has the area $(3)^{1/2}a^2/4$. Each atom is bonded to three atoms below and six in the surface, so there are three dangling bonds, for a dangling bond density of $4(3)^{1/2}/a^2$, and a surface energy density γ of $4(3)^{1/2}\varepsilon/a^2$. These results are gathered together in Table 1.2.

1.2.4. Surfaces of Zinc Blende and Diamond Structures

The remaining structure that is of interest is the diamond structure depicted in Fig. 1.11. This consists of a set of fcc atoms together with a second set occupying half of the tetrahedral holes of the first set, as was discussed above. We will discuss surfaces of ZnS crystals in which the small zinc atoms reside in the tetrahedral

TABLE 1.2. Atom Density, Number of Bonds Broken per Atom, and Surface Energy Density for Various Facets or Surfaces of FCC centered cubic and diamond structures[a]

Structure and Facet	Atom Density	Bonds per Atom	Surface-energy Density
Fcc [100]	$2/a^2$	4	$8\varepsilon/a^2$
Fcc [110]	$(2)^{1/2}/a^2$	5	$5(2)^{1/2}\ \varepsilon/a^2$
Fcc [111]	$4/a^2$	3	$4(3)^{1/2}\ \varepsilon/a^2$
Diamond [100]	$2/a^2$	2	$4\varepsilon/a^2$
Diamond [110]	$2(2)^{1/2}/\mathrm{a}^2$	1	$2(2)^{1/2}\ \varepsilon/a^2$
Diamond [111]	$4/a^2$	1	$4\varepsilon/(3)^{1/2}a^2$

[a] *Key*: $a=$ lattice constant; $\varepsilon =$ surface energy per dangling chemical bond.

sites of the dominant fcc sulfur lattice. The atom locations in various faces can be deduced from Figs. 1.11b and 1.12.

Figure 1.17, with the smaller octahedral site atoms omitted, depicts the [001] surface, which corresponds to the top face of the cubic unit cell in Figs. 1.11b and 1.12. The atom density of this face is $2/a^2$, the same as in the simple fcc case discussed above. It is clear from Fig. 1.11b that each surface atom has two dangling bonds, so the dangling bond density is $4/a^2$. Figure 1.18 sketches the [110] surface, which contains four large edge and apex S atoms of the surface unit cell, plus two smaller interior Zn atoms depicted by small open circles in the lower part of Fig. 1.18. In this figure the medium-sized octahedral site atoms and the shaded small tetrahedral site atoms are absent, so there are four atoms per surface unit cell of area $(2)^{1/2}a^2$. It is clear from Fig. 1.11b that each atom has two bonds within the surface and one below it, so there is one dangling bond per atom, for a dangling bond density of $2(2)^{1/2}/a^2$. For simplicity we presume a surface energy density $2(2)^{1/2}a^2$, assuming that the dangling bond energy ε_i of the zinc and sulfur atoms has the same value ε, which is really not the case. Figure 1.14 depicts the atoms associated with the [111] plane. The unshaded circles represent the atoms on the surface, and the shaded circles correspond to the first layer below the surface. Either type could be sulfur, and the other one would be zinc. The atom density is $4/(3)^{1/2}a^2$ as in the fcc cubic case discussed above. The dots centered in the open circles denote dangling bonds, one such bond per atom, directed upward perpendicular to the surface. Hence the dangling bond density is $4/(3)^{1/2}a^2$.

We showed above how the energy of the atoms close to the [111] face of a diamond or zinc blende crystal can decrease by a reconstruction in which the atoms rearrange to a lattice with a bimolecular unit cell formed from the pairs of atoms depicted in Fig. 1.15. These surface atoms bond in pairs so that there are no dangling bonds remain and the energy is lowered. In some cases the spacing between layers slightly varies between the surface and the bulk, with the layers near the surface undergoing an expansion outward or a contraction inward. In Fig. 1.14 the atoms with upward-directed bonds formed the surface. If the layer below had formed the surface, then each surface atom would have had three dangling bonds, a higher-energy situation, which is therefore unlikely to occur. Surfaces tend to form which minimize the density of dangling bonds.

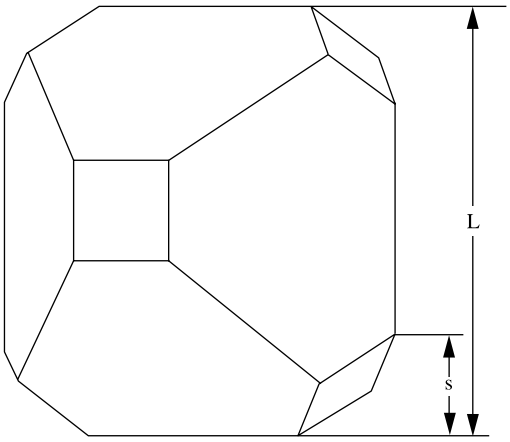

Figure 1.20. Truncated octahedron. For a fcc crystal the atoms on the square faces have a square lattice configuration, and the atoms on the hexagon shaped faces have a hexagonal close packed coniiguration. (From G. Burns, *Solid State Physics*, Academic Press, Orlando, FL, 1985, p. 676.)

Sometimes silicon forms a crystal with the external shape of an octahedron which has eight [111] faces that minimize the surface energy. On other occasions the octahedron is truncated in the manner illustrated in Fig. 1.20 since a truncated configuration has a lower surface energy for the same volume of crystal.

A favorable technique for studying surface structure is low-energy electron diffraction (LEED). Another technique is reflection high-energy electron diffraction (RHEED) carried out at grazing incidence. These techniques are discussed in Chapter 2.

1.2.5. Adsorption of Gases

Since surfaces generally have dangling bonds, they can be chemically very reactive. Owing to this reactivity, they readily adsorb gases in contact with them. Adsorbed molecules are held in place on the surface by a binding energy ΔH_{ads} called the *heat of adsorption*. When chemical bonding takes place between surface atoms and adsorbed atoms or molecules, the phenomenon is called *chemisorption*. This bonding can be of either a covalent or an ionic type. If a much weaker van der Waals type of bonding occurs the phenomenon is referred to as *physisorption*. It is customary to call the process *physisorption* when the heat of adsorption is less than $\sim$10 kcal/mol, and to label it *chemisorption* when the heat of adsorption exceeds 10 kcal/mol. For example, H_2 and O_2 molecules chemisorb on molybdenum with ΔH_{ads} values of 40 kcal/mol and 172 kcal/mol, respectively.

The surface of a solid can be considered as the outermost layer of atoms plus the region between 0.5 and 1.5 nm above and below it. This includes the bonding electrons on the inside and the dangling bonds on the outside. We are talking about an

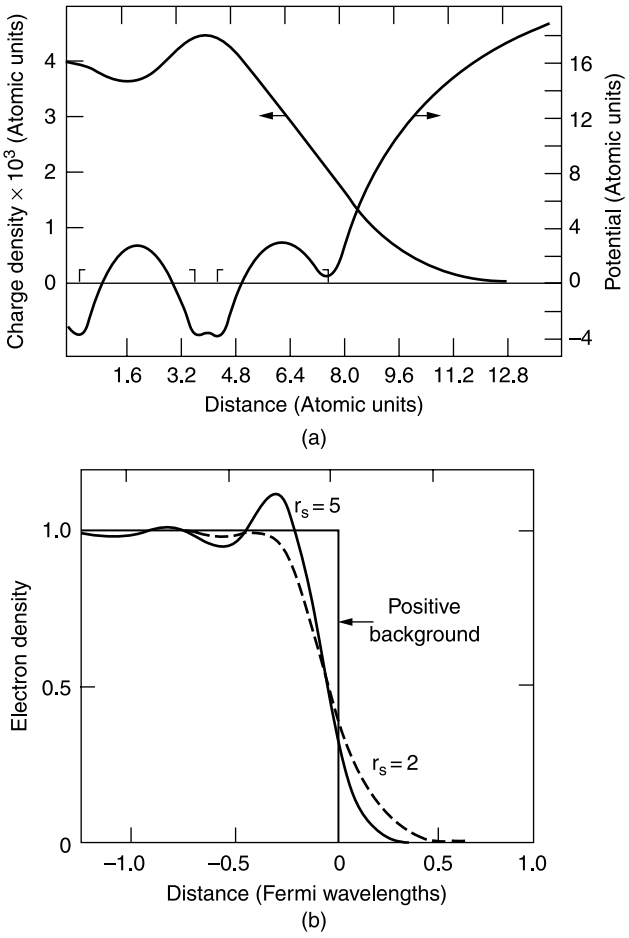

Figure 1.21. (a) Calculated potential energy and charge density as a function of z for metallic sodium. The Na^+ ion cores of the last two atomic layers are indicated by brackets. [From J. A. Applebaum and D. R. Hamann, *Phys. Rev.* **B6**, 2166 (1972)]. (b) Calculated electron density near the surface of metallic Na, where $r_s = 2$ Å is close to the Na value $r_s = 2.08$ Å. [From N. D. Lang and W. Kohn, *Phys. Rev.* **B1**, 4555 (1970). See also G. Burns (cited in Fig. 1.20 legend), p. 703.]

atomically clean surface, one exposed to an ultrahigh vacuum, and which contains no adsorbed atoms or molecules. An ultrahigh vacuum is one in which the pressure is less than 10^{-8} Pa ($<7.5 \times 10^{-11}$ torr). When a molecule from the surrounding gas strikes the surface, there is a probability S called the *sticking coefficient* that it will stick to the surface, that is, that it will be adsorbed. In a standard vacuum of 10^{-6} torr it will take surrounding gas molecules with a sticking coefficient $S = 1$ about one second to build up a monolayer, whereas at 10^{-10} torr the monolayer buildup time is much longer than an hour.

1.2.6. Electronic Structure of a Surface

The formation of a surface by cleaving a crystal in an ultrahigh vacuum can lead to a rearrangement of the electrons on the atoms that constitute the surface layer. We will discuss the case of a metallic surface. A double layer of charge forms because the center of gravity of the conduction electron charge differs from the center of gravity of the atomic nuclei. The charge density and potential energy at a [001] surface of sodium, which is fcc, vary with the distance perpendicular to the surface in the manner shown in Fig. 1.21a. The calculated electron density varies with this distance in the manner shown in Fig. 1.21b. The calculations were carried out for two values $r_s = 2$ Å and $r_s = 5$ Å of the equivalent electron radius (radius of average spherical volume occupied by a conduction electron), where r_s equals 2.08 Å and 1.41 Å for Na and Cu, respectively.

An important surface-related property of a metal is its workfunction φ, which is the energy required to remove an electron from the surface and place it an infinite distance away. The workfunction for polycrystalline sodium is 2.75 eV. When sodium is chemisorbed on the surface of another metal the workfunction varies with the surface coverage in the manner shown on Fig. 1.22. The figure shows calculations carried for sodium and for cesium chemisorbed on an aluminum substrate. As the coverage increases, the workfunction initially drops from its initial value for a bare aluminum surface, reaches a minimum, and then rises to the value of φ corresponding to that of the covering metal Na or Cs. The observed effect occurs because the adsorbed alkali atom Na loses an electron to the aluminum and becomes a positive ion or cation Na^+. The result is the creation of a dipole double layer with a polarity opposite that of the initial bare aluminum surface, so the magnitude of the net dipole moment density decreases, and φ also decreases.

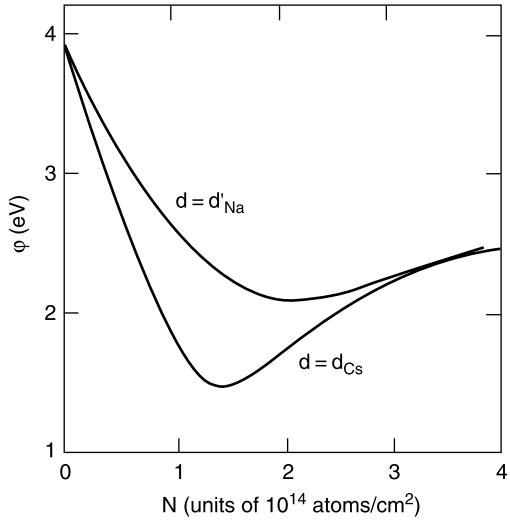

Figure 1.22. Calculated workfunction versus Na and Cs surface coverage for $r_s = 2$ Å, which is close to the value for aluminum. [From N. D. Lang, *Phys. Rev.* **B4**, 4234 (1971).]

1.2.7. Surface Quantum Well

A *quantum well* is a three-dimensional structure in which two dimensions are large, perhaps a centimeter or a millimeter, and the third is in the nanometer range. The surface of a solid, with or without the presence of a layer of adsorbed gas molecules, may be regarded as a quantum well. It has a regular structure, its lateral dimensions are macroscopic, and its thickness is several nanometers. It has characteristic charge and potential energy distributions, as discussed above. Quantum wells are discussed in Chapter 9.

1.3. ENERGY BANDS

1.3.1. Insulators, Semiconductors, and Conductors

When a solid is formed, the energy levels of the atoms broaden and form bands with forbidden gaps between them. The electrons can have energy values that exist within one of the bands, but cannot have energies corresponding to values in the gaps between the bands. The lower-energy bands due to the inner atomic levels are narrower and are all full of electrons, so they do not contribute to the electronic properties of a material. They are not shown in the figures. The outer or valence electrons that bind the crystal together occupy a *valence band*, which, for an insulating material, is full of electrons that cannot move since they are fixed in position in chemical bonds. There are no delocalized electrons to carry current, so the material is an insulator. The conduction band is far above the valence band in energy, as shown in Fig. 1.23a, so it is not thermally accessible, and remains essentially empty. In other words, the heat content of the insulating material at room temperature $T = 300$ K is not sufficient to raise an appreciable number of electrons from the valence band to the conduction band, so the number in the conduction band is negligible. Another way to express this is to say that the value of the gap energy E_g far exceeds the value $k_B T$ of the thermal energy, where k_B is Boltzmann's constant.

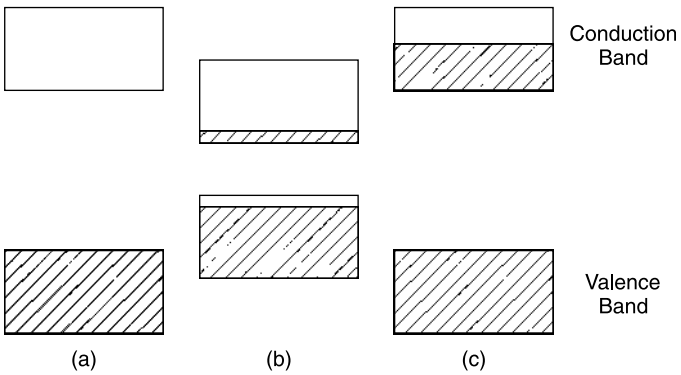

Figure 1.23. Energy bands of (a) an insulator, (b) an intrinsic semiconductor, and (c) a conductor. The crosshatching indicates the presence of electrons in the bands.

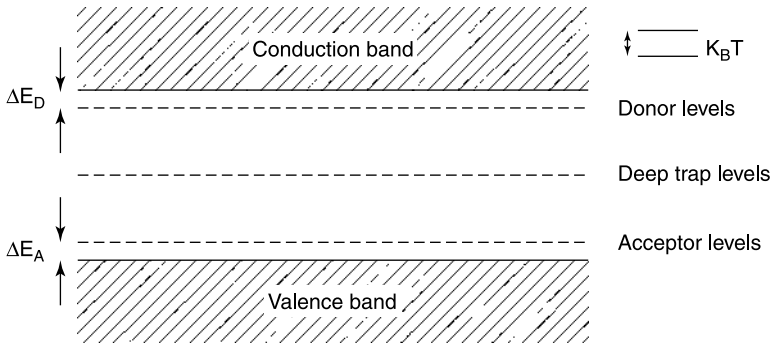

Figure 1.24. Sketch of the forbidden energy gap showing acceptor levels the typical distance Δ_A above the top of the valence band, donor levels the typical distance Δ_D below the bottom of the conduction band, and deep-trap levels nearer to the center of the gap. The value of the thermal energy k_BT is indicated on the right.

For a semiconductor, the gap between the valence and conduction bands is much less, as shown in Fig. 1.23b, so E_g is closer to the thermal energy k_BT, and the heat content of the material at room temperature can bring about the thermal excitation of some electrons from the valence band to the conduction band where they carry current. The density of electrons reaching the conduction band by this thermal excitation process is relatively low, but by no means negligible, so the electrical conductivity is small; hence the name *semiconducting*. A material of this type is called an *intrinsic semiconductor*. A semiconductor can be doped with donor atoms that donate electrons to the conduction band where they can carry current. The material can also be doped with acceptor atoms that obtain electrons from the valence band and leave behind positive charges called *holes* which can also carry current. The energy levels of these donors and acceptors lie in the energy gap, as shown in Fig. 1.24. The former produces N type, that is, negative charge or electron conductivity, and the latter produces P type, that is, positive charge or hole conductivity, as will be clarified in Section 1.4.1. These two types of conductivity in semiconductors are temperature dependent, as is the intrinsic semiconductivity.

A *conductor* is a material with a full valence band, and a conduction band partly full with delocalized conduction electrons that are efficient in carrying electric current. The positively charged metal ions at the lattice sites have given up their electrons to the conduction band, and constitute a background of positive charge for the delocalized electrons. Figure 1.23c shows the energy bands for this case. In actual crystals the energy bands are much more complicated than is suggested by the sketches of Fig. 1.23, with the bands depending on the direction in the lattice, as we shall see below.

1.3.2. Reciprocal Space

In Section 1.1.2 we discussed the structures of different types of crystals in ordinary or coordinate space. These provided us with the positions of the atoms in the lattice. To treat the motion of conduction electrons, it is necessary to consider a different type

of space, which is mathematically called a *dual space* relative to the coordinate space. This dual or reciprocal space arises in quantum mechanics, and a brief qualitative description of this space is presented here.

The basic relationship between the frequency $f = \omega/2\pi$, the wavelength λ, and the velocity v of a wave is $\lambda f = v$. It is convenient to define the wavevector $k = 2\pi/\lambda$ to give $f = (k/2\pi)v$. For a matter wave, or the wave associated with conduction electrons, the momentum $p = mv$ of an electron of mass m is given by $p = (h/2\pi)k$, where Planck's constant h is a universal constant of physics. Sometimes a reduced Planck's constant $\hbar = h/2\pi$ is used, where $p = \hbar k$. Thus, for this simple case the momentum is proportional to the wavevector k, and k is inversely proportional to the wavelength with the units of reciprocal length, or reciprocal meters. We can define a reciprocal space called *k-space* to describe the motion of electrons.

If a one-dimensional crystal has a lattice constant a and a length that we take to be $L = 10a$, then the atoms will be present along a line at positions $x = 0,a,2a,3a, \ldots, 10a = L$. The corresponding wavevector k will assume the values $k = 2\pi/L, 4\pi/L, 6\pi/L, \ldots, 20\pi/L = 2\pi/a$. We see that the smallest value of k is $2\pi/L$ and the largest value is $2\pi/a$. The unit cell in this one-dimensional coordinate space has length a, and the important characteristic cell in reciprocal space, called the *Brillouin zone*, has the value $2\pi/a$. The electron sites within the Brillouin zone are at the reciprocal lattice points $k = 2\pi n/L$, where for our example $n = 1,2,3,\ldots,10$, and $k = 2\pi/a$ at the Brillouin zone boundary where $n = 10$.

For a rectangular direct lattice in two dimensions with coordinates x and y, and lattice constants a and b, the reciprocal space is also two-dimensional with the wavevectors k_x and k_y. By analogy with the direct lattice case, the Brillouin zone in this two-dimensional reciprocal space has length $2\pi/a$ and width $2\pi/b$, as sketched in Fig. 1.25. The extension to three dimensions is straightforward. It is important to keep in mind that k_x is proportional to the momentum p_x of the conduction electron in the x direction, and similarly for the relationship between k_y and p_y.

1.3.3. Energy Bands and Gaps of Semiconductors

The electrical, optical, and other properties of semiconductors depend strongly on how the energy of the delocalized electrons involves the wavevector k in reciprocal or k-space, with the electron momentum p given by $p = mv = \hbar k$, as explained

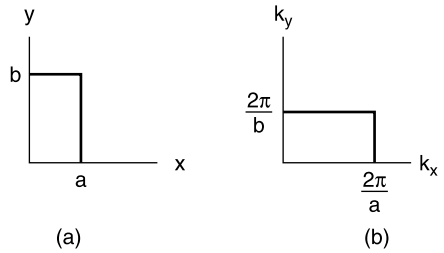

Figure 1.25. Sketch of (a) unit cell in two-dimensional x,y coordinate space; (b) corresponding Brillouin zone in k_x,k_y reciprocal space for a rectangular Bravais lattice.

above. We will consider three-dimensional crystals; in particular, we are interested in the properties of the III–V and the II–VI semiconducting compounds that have a cubic structure, so their three lattice constants are the same, namely, $a = b = c$. The electron motion expressed in the coordinates k_x,k_y,k_z of reciprocal space takes place in the Brillouin zone, and the shape of this zone for these cubic compounds is shown in Fig. 1.26. Points of high symmetry in the Brillouin zone are designated by capital (uppercase) Greek or Roman letters, as indicated.

The energy bands depend on the position in the Brillouin zone, and Fig. 1.27 presents these bands for the intrinsic (i.e., undoped) III–V compound GaAs. The figure plots energy versus the wavevector k in the following Brillouin zone directions: along Δ from point Γ to X, along Λ from Γ to L, along Σ from Γ to K, and along the path between points X and K. These points and paths are indicated in the sketch of the Brillouin zone in Fig. 1.30. We see from Fig. 1.27 that the various bands have prominent maxima and minima at the centerpoint Γ of the Brillouin zone. The energy gap or region where no band appears extends from the zero of energy at point Γ_8 to point Γ_6 directly above the gap at the energy $E_g = 1.35$ eV. The bands below point Γ_8 constitute the valence band, and those above point Γ_6 form the conduction band. Hence Γ_6 is the lowest-energy point of the conduction band, and Γ_8 is the highest point of the valence band.

At absolute zero all of the energy bands below the gap are filled with electrons, and all the bands above the gap are empty, so at $T = 0$ K the material is an insulator. At room temperature the gap is sufficiently small so that some electrons are thermally excited from the valence band to the conduction band, and these relatively few excited electrons gather in the region of the conduction band immediately above its minimum at Γ_6, a region that is referred to as a "valley." These electrons carry some electric current; hence the material is a semiconductor. Gallium arsenide is called a *direct-bandgap semiconductor* because the top of the valence band and the bottom of the conduction band are both at the same centerpoint (Γ) in the

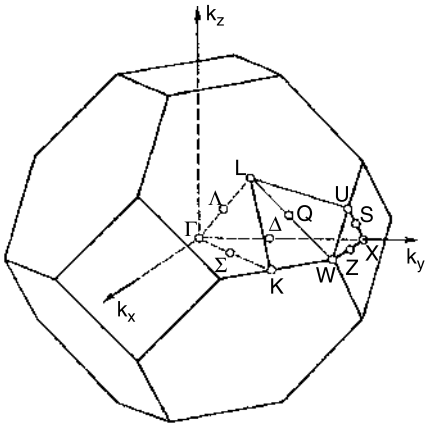

Figure 1.26. Brillouin zone of the gallium arsenide and zinc blende semiconductors showing the high-symmetry points Γ, K, *L*, *U*, *W*, and *X*, and the high-symmetry lines Δ, Λ, Σ, *Q*, *S*, and *Z*. (From G. Burns, *Solid State Physics*, Academic Press, Boston, 1985, p. 302.)

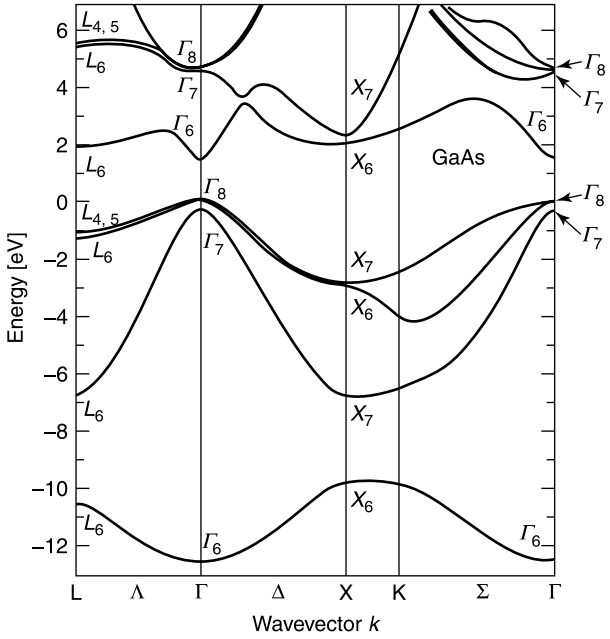

Figure 1.27. Band structure of the semiconductor GaAs calculated by the pseudopotential method. (From M. L. Cohen and J. Chelikowsky, *Electronic Structure and Electronic Properties of Semiconductors*, 2nd ed., Springer-Verlag, Solid State Science series, Vol. 75, Springer, Berlin, 1989, p. 65.)

Brillouin zone, as is clear from Fig. 1.27. Electrons in the valence band at point Γ_8 can become thermally excited to point Γ_6 in the conduction band with no change in the wavevector k. The compounds GaAs, GaSb, InP, InAs, and InSb and all the II–VI compounds included in Table B.6 have direct gaps. In some semiconductors such as Si and Ge the position of the top of the valence band in the Brillouin zone differs from the bottom of the conduction band, and these are called *indirect-gap semiconductors*.

Figure 1.28 depicts the situation at point Γ of a direct-gap semiconductor on an expanded scale, at temperatures above absolute zero, with energy bands approximated by parabolas. The conduction band valley at Γ_6 is shown occupied by electrons up to the Fermi level, which is defined as the energy of the highest occupied state. The excited electrons leave behind empty states near the top of the valence band, and these act like positive charges called *holes* in an otherwise full valence band. These hole levels exist above the energy $-E_F'$, as indicated in Fig. 1.28. Since an intrinsic or undoped semiconductor has just as many holes in the valence band as it has electrons in the conduction band, the corresponding volumes filled with these electrons and holes in k-space are equal to each other. These electrons and holes are the charge carriers of current, and the temperature dependence of their concentrations in GaAs, Si, and Ge is given in Fig. 1.29.

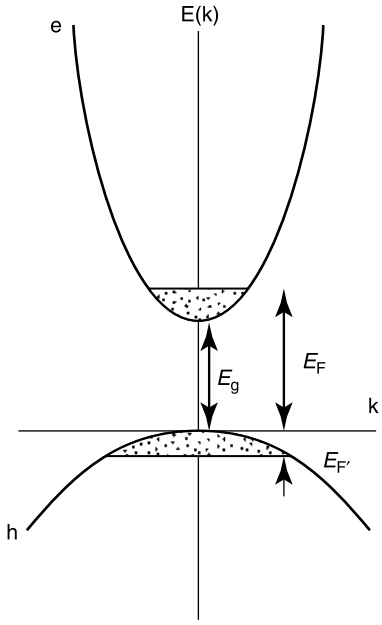

Figure 1.28. Sketch of the lower valence band and the upper conduction band of a semiconductor approximated by parabolas. The region in the valence band containing holes and that in the conduction band containing electrons are crosshatched. The Fermi energies E_F and $E_F{}'$ mark the highest occupied level of the conduction band and the lowest unoccupied level of the valence band, respectively. The zero of energy is taken as the top of the valence band, and the direct-bandgap energy E_g is indicated.

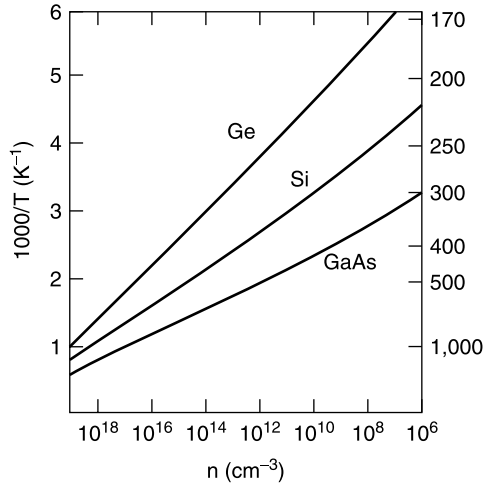

Figure 1.29. Temperature dependencies of the intrinsic carrier density of the semiconductors Ge, Si, and GaAs. (From G. Burns, *Solid State Physics*, Academic Press, Boston, 1985, p. 315.)

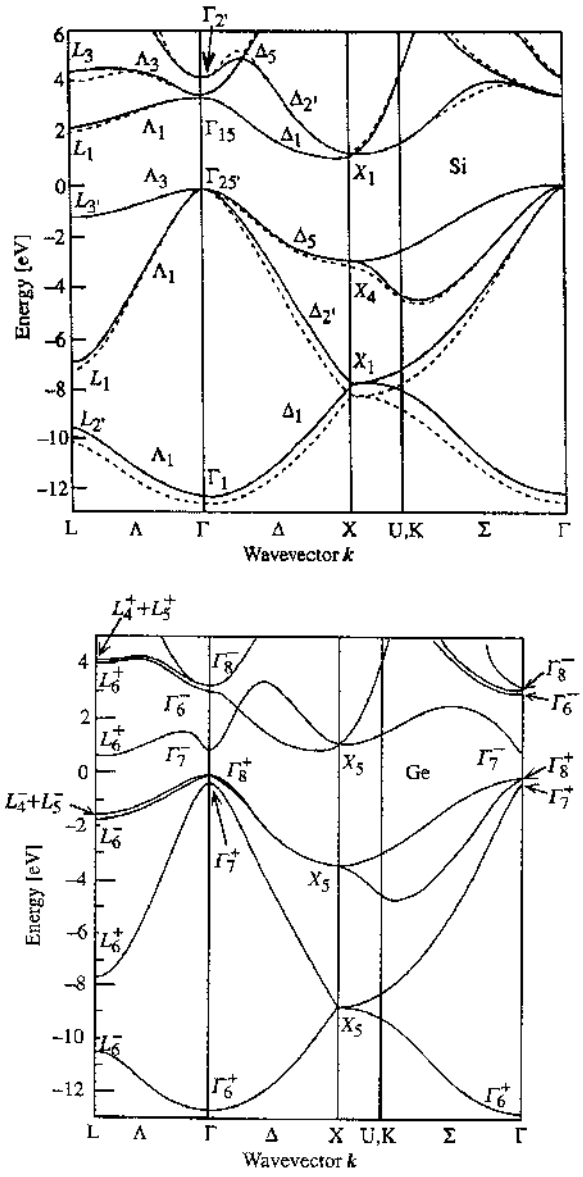

Figure 1.30. Band structure plots of the energy bands of the indirect-gap semiconductors Si and Ge. The bandgap (absence of bands) lies slightly above the Fermi energy $E = 0$ on both figures, with the conduction band above and the valence band below the gap. The figure shows that the lowest point or bottom of the conduction band of Ge is at the energy $E = 0.6$ at the symmetry point L (labeled L_6^+), and for Si it is 85% of the way along the direction Δ_1 from Γ_{15} to X_1. It is clear from Fig. 1.27 that the bottom of the conduction band of the direct-gap semiconductor GaAs is at the symmetry point Γ_6, The top of the valence band is at the centerpoint Γ of the Brillouin zone for all three materials. (From M. L. Cohen and J. Chelikowsky, *Electronic Structure and Electronic Properties of Semiconductors*, 2nd ed., Springer-Verlag, Solid State Science series, Vol. 75, Springer, Berlin, 1989, pp. 53, 64.)

In every semiconductor listed in Table B.6, including Si and Ge, the top of its valence band is located at the center of the Brillouin zone, but in the indirect-bandgap semiconductors Si, Ge, AlAs, AlSb, and GaP the lowest valley of their conduction bands appears at different locations in k-space and at point Γ. This is shown in Fig. 1.30 for the indirect-bandgap materials Si and Ge. We see from Fig. 1.30 that the conduction band of Ge is minimum at point L, which is in the middle of the hexagonal face of the Brillouin zone along the Λ or [111] direction depicted in Fig. 1.26. The Brillouin zone has eight such faces, with each point L shared by two zones, so the zone actually contains only four of these points proper to it, and we say that the valley degeneracy for Ge is 4. The semiconductor Si exhibits a lowest conduction band minimum along the Δ or [001] direction about 85% of the way to point X, as shown in Fig. 1.30. The corresponding valley of the compound GaP, not shown, also is located along Δ about 92% of the way to X. We see from Fig. 1.26 that there are six such Δ lines in the Brillouin zone, so this valley degeneracy is 6. Associated with each of the valleys that we have been discussing, at point L for Ge and along direction Δ for Si, there is a three-dimensional constant-energy surface in the shape of an ellipsoid that encloses the conduction electrons in the corresponding valleys, and these ellipsoids are sketched in Figs. 1.31a and 1.31b for Ge and Si, respectively.

Some interesting experiments such as cyclotron resonance have been carried out to map the configuration of these ellipsoid-type constant-energy surfaces. In a cyclotron resonance experiment conduction electrons are induced to move along

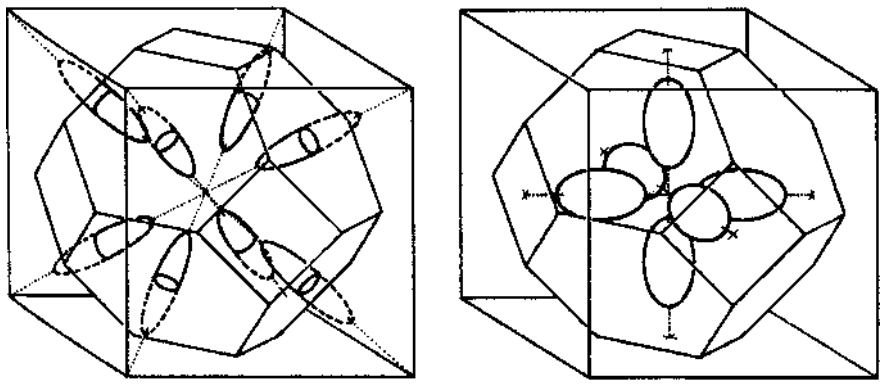

Figure 1.31. Ellipsoidal constant energy surfaces in the conduction band of germanium (a) and silicon (b). The constant-energy surfaces of Ge are aligned along symmetry direction Λ and centered at symmetry point L. As a result, they lie half inside (solid lines) and half outside (dashed lines) the first Brillouin zone, so this zone contains the equivalent of four complete energy surfaces. The surfaces of Si lie along the six symmetry directions Δ (i.e., along $\pm k_x$, $\pm k_y$, $\pm k_z$), and are centered 85% of the way from the centerpoint Γ to symmetry point X. All six of them lie entirely within the Brillouin zone, as shown. Figure 1.26 shows the positions of symmetry points Γ, L, and X, and of symmetry lines Δ and Λ, in the Brillouin zone. (From G. Burns, *Solid State Physics*, Academic Press, Boston, 1985, p. 313.)

constant-energy surfaces at a velocity that always remains perpendicular to an applied magnetic field direction. By utilizing various orientations of applied magnetic fields relative to the ellipsoid, the electrons at the surface execute a variety of orbits, and by measuring the trajectories of these orbits, we can delineate the shape of the energy surface.

Table B.6 lists values of the energy gap E_g for types III–V and II–VI semiconductors at room temperature (left-hand value) and in several cases it also lists the value at absolute zero temperature (right-hand value). Table B.7 gives the temperature and pressure dependences, dE_g/dT and dE_g/dP, respectively, of the gap at room temperature.

1.3.4. Effective Mass

The kinetic energy of a conduction electron moving at a velocity v in a semiconductor is found to have the value $\frac{1}{2}m^*v^2 = p^2/2m^*$, where $\mathbf{p} = m^*\mathbf{v}$ is the momentum, and often the mass m^* in these expressions differs from the electron's fundamental or rest mass value m_0, which is a universal constant, so it is called an *effective mass*. This is analogous to the case of the kinetic energy of an object of mass m moving at a velocity v through a liquid. Its effective mass is reduced by the volume of the displaced liquid, and hence its measurable kinetic energy has a value lower than $\frac{1}{2}mv^2$.

On a simple one-dimensional model the energy E of a conduction electron has a quadratic dependence on the wavevector $\mathbf{k} = \mathbf{p}/h$ through the expression

$$E = \hbar^2 k^2/2m^* \tag{1.12}$$

The first derivative of this expression provides the velocity v

$$\frac{1}{\hbar}\frac{dE}{dk} = \frac{hk}{m^*} = v \tag{1.13}$$

and the second derivative provides the effective mass m^*

$$\frac{1}{\hbar^2}\frac{d^2E}{dk^2} = \frac{1}{m^*} \tag{1.14}$$

which, as we said above, differs in general from the free-electron mass m_0. These equations are rather trivial for the simple parabolic energy expression (1.12), but we see from the energy bands of Figs. 1.27 and 1.30 that the actual dependence of the energy E on k is much more complex than Eq. (1.12) indicates. This equation (1.14) is a general definition of the effective mass m^*, and the wavevector k dependence of m^* can be evaluated from the band structure plots by carrying out the differentiations. We see from a comparison of the slopes near the conduction band minimum and the valence band maximum of GaAs at point Γ on Fig. 1.27

(see also Fig. 1.28) that the upper electron bands have steeper slopes and hence lighter masses than do the lower hole bands with more gradual slopes.

1.3.5. Fermi Surfaces

At very low temperatures electrons fill the energy bands of solids up to the Fermi energy E_F, and the bands are empty for energies that exceed E_F. In three-dimensional k-space the set of values of k_x,k_y,k_z that satisfy the equation $\hbar^2(k_x^2 + k_y^2 + k_z^2)/2m = E_F$ form the *Fermi surface*. All k_x,k_y,k_z energy states that lie below this surface are full, and the states above the surface are empty. The Fermi surface encloses all the electrons in the conduction band that carry electric current. In the good conductors copper and silver the conduction electron density is 8.5×10^{22} and 5.86×10^{22} electrons/cm^3, respectively. From another viewpoint, the Fermi surface of a good conductor can fill the entire Brillouin zone. In the intrinsic semiconductors GaAs, Si, and Ge the carrier density at room temperature from Fig. 1.29 is approximately 10^6, 10^{10}, and 10^{13} carriers/cm^3, respectively, many orders of magnitude below that of metals, and semiconductors are seldom doped to concentrations above 10^{19} centers/cm^3. An *intrinsic semiconductor* is one with a full valence band and an empty conduction band at absolute zero of temperature. As we saw above, at ambient temperatures some electrons are thermally excited to the bottom of the conduction band, and an equal number of empty sites or holes are left behind near the top of the valence band. This means that only a small percentage of the Brillouin zone contains electrons in the conduction band, and the number of holes in the valence band is correspondingly small. In a one-dimensional representation this reflects the electron and hole occupancies depicted in Fig. 1.28.

If the conduction band minimum is at point Γ in the center of the Brillouin zone, as is the case with GaAs, then the Fermi surface will be very close to a sphere since the symmetry is cubic, and to a good approximation we can assume a quadratic dependence of the energy on the wavevector k, corresponding to Eq. (1.12). Therefore the Fermi surface in k-space is a small sphere given by the standard equation for a sphere

$$E_F = E_g + \frac{\hbar^2}{2m_e}(k_x^2 + k_y^2 + k_z^2)_F = E_g + \frac{\hbar^2}{2m_e}k_F^2 \tag{1.15}$$

where E_F is the Fermi energy and the electron mass m_e relative to the free-electron mass has the values given in Table B.8 for various direct-gap semiconductors. Equations (1.15) are valid for direct-gap materials where the conduction band minimum is at point Γ. In all the semiconductor materials under consideration the valence band maxima are located at the Brillouin zone centerpoint Γ.

The situation for the conduction electron Fermi surface is more complicated for indirect gap semiconductors. We mentioned above that Si and GaP exhibit conduction band minima along the Δ direction of the Brillouin zone, and the corresponding six ellipsoidal energy surfaces are sketched in Fig. 1.31a. The longitudinal and

transverse effective masses, m_L and m_T, respectively, have the following values

$$\frac{m_L}{m_0} = 0.92 \quad \frac{m_T}{m_0} = 0.19 \quad \text{for} \quad \text{Si} \tag{1.16a}$$

$$\frac{m_L}{m_0} = 7.25 \quad \frac{m_T}{m_0} = 0.21 \quad \text{for} \quad \text{GaP} \tag{1.16b}$$

for these two indirect-gap semiconductors. Germanium has its band minimum at points L of the Brillouin zone sketched in Fig. 1.26, and its Fermi surface is the set of ellipsoids centered at paints L with their axes along the Λ or [111] directions, as shown in Fig. 1.31a. The longitudinal and transverse effective masses for these ellipsoids in Ge are $m_L/m_0 = 1.58$ and $m_T/m_0 = 0.081$, respectively. Cyclotron resonance techniques together with the application of stress to the samples can be used to determine these effective masses for the indirect-bandgap semiconductors.

1.4. LOCALIZED PARTICLES

1.4.1. Donors, Acceptors, and Deep Traps

When an atom in column V of the periodic table, such as P, As, or Sb, which has five electrons in its outer or valence electron shell, is a substitutional impurity in Si, it uses four of these electrons to satisfy the valence requirements of the four nearest-neighbor silicons, and the one remaining electron remains weakly bound. The atom easily donates or passes on this electron to the conduction band, so it is called a *donor*, and the electron is called a *donor electron*. This occurs because the donor energy levels lie in the forbidden region close to the conduction band edge by the amount ΔE_d relative to the thermal energy value $k_B T$, as indicated in Fig. 1.24. A Si atom substituting for Ga plays the role of a donor in GaAs, Al substituting for Zn in ZnSe serves as a donor, and so on. The conductivity that is associated with the flow of electric current carried by these conduction band electrons is called *N-type*.

An atom from column III atom such as Al or Ga, called an *acceptor atom*, which has three electrons in its valence shell can serve as a substitutional defect in Si, and in this role it requires four valence electrons to bond with the tetrahedron of nearest-neighbor Si atoms. To accomplish this it draws or accepts an electron from the valence band, leaving behind a hole at the top of this band. The conductivity corresponding to electric current carried by the movement of these holes is called *P-type*. The transfer of an electron to Si occurs easily because the energy levels of the acceptor atoms are in the forbidden gap slightly above the valence band edge by the amount ΔE_A, which is small relative to $k_B T$, as indicated in Fig. 1.24. In other words, the excitation energies needed to ionize the donors and to add electrons to the acceptors are much less than the thermal energy at room temperature $T = 300$ K (i.e., $\Delta E_D, \Delta E_A \ll k_B T$), so virtually all donors are positively ionized and virtually all acceptors are negatively ionized at room temperature.

The donor and acceptor atoms that we have been discussing are known as "shallow centers," that is, shallow traps of electrons or holes, because their excitation energies are much less than that of the bandgap (ΔE_D, $\Delta E_A \ll E_g$). There are other centers with energy levels that lie deep within the forbidden gap, often closer to its center than to the top or bottom, in contrast to the case with shallow donors and acceptors. Since generally $E_g \gg k_B T$, these traps are not extensively ionized, and the energies involved in exciting or ionizing them are not small. Examples of deep centers are defects associated with broken bonds, or strain involving displacements of atoms.

1.4.2. Mobility

Another important parameter of a semiconductor is the mobility μ or charge carrier drift velocity v per unit electric field E, given by the expression $\mu = |v|/E$. This parameter is defined as positive for both electrons and holes. Table B.9 lists the mobilities μ_e and μ_h for electrons and holes, respectively, in the semiconductors under consideration. The electrical conductivity σ is the sum of contributions from the concentrations of electrons n and of holes p in accordance with the expression

$$\sigma = (ne\mu_e + pe\mu_h) \tag{1.17}$$

where e is the electronic charge. The mobilities have a weak power-law temperature dependence T^n, and the pronounced T dependence of the conductivity is due principally to the dependence of the electron and hole concentrations on the temperature. In doped semiconductors this generally arises mainly from the Boltzmann factor $\exp(-E_i/k_B T)$ associated with the ionization energies E_i of the donors or acceptors. Typical ionization energies for donors and acceptors in Si and Ge listed in Table B.10 range from 0.0096 to 0.16 eV, which is much less than the bandgap energies 1.12 eV and 0.67 eV of Si and Ge, respectively. Figure 1.24 shows the locations of donor and acceptor levels on an energy band plot, and makes clear that their respective ionization energies are much less than E_g. The thermal energy $k_B T = 0.026\,\text{eV}$ at room temperature ($T = 300\,\text{K}$) is often comparable to the ionization energies. In intrinsic or undoped materials the main contribution is from the exponential factor $\exp(-E_g/2\,k_B T)$ in the following expression from the law of mass action

$$n_i = p_i = 2K\left(\frac{k_B T}{2\pi\hbar^2}\right)^{3/2} (m_e m_h)^{3/4} \exp\frac{-E_g}{2k_B T} \tag{1.18}$$

where K is a constant and the intrinsic concentrations of electrons n_i and holes p_i are equal to each other because the thermal excitation of n_i electrons to the conduction band leaves behind the same number p_i of holes in the valence band, that is, $n_i = p_i$. We see that the expression (1.18) contains the product $m_e m_h$ of the effective masses m_e and m_h of the electrons and holes, respectively, and the ratios m_e/m_0 and m_h/m_0 of these effective masses to the free-electron mass m_0 are presented in

Table B.8. These effective masses strongly influence the properties of excitons to be discussed next.

1.4.3. Excitons

An ordinary negative electron and a positive electron, called a *positron*, situated a distance r apart in free space experience an attractive *Coulomb force*, which has the value $-e^2/4\pi\varepsilon_0 r^2$, where e is their charge and ε_0 is the dielectric constant of free space. A quantum-mechanical calculation shows that the electron and positron interact to form an atom called *positronium*, which has bound state energies given by the Rydberg formula introduced by Niels Bohr in 1913 to explain the hydrogen atom

$$\mathrm{E} = -\frac{e^2}{8\pi\varepsilon_0 a_0 n^2} = -\frac{6.8}{n^2}\mathrm{eV} \tag{1.19}$$

where a_0 is the Bohr radius given by $a_0 = 4\pi\varepsilon_0\hbar^2/m_0e^2 = 0.0529\,\mathrm{nm}$, m_0 is the free-electron (and positron) mass, and the quantum number n takes on the values $n = 1,2,3,\ldots,4$. For the lowest-energy or ground state, which has $n = 1$, the energy is 6.8 V, which is exactly half the ground-state energy of a hydrogen atom, since the effective mass of the bound electron–positron pair is half that of the bound electron–proton pair in the hydrogen atom. Figure 1.32 shows the energy

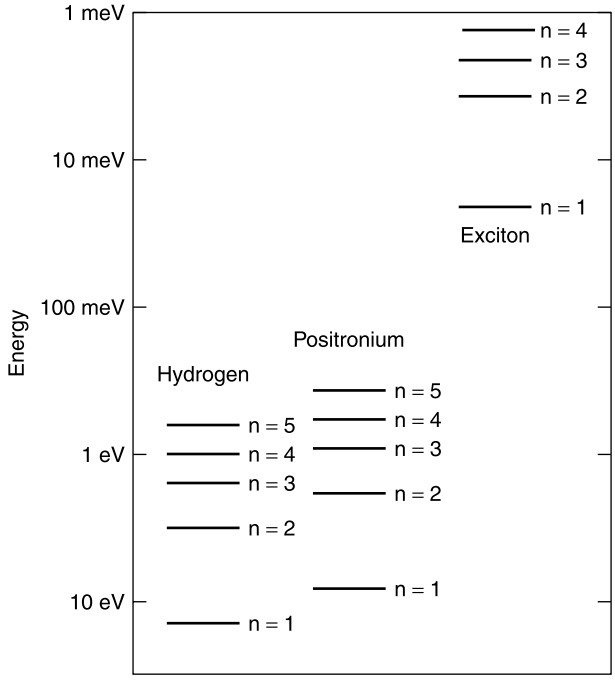

Figure 1.32. The first few energy levels in the Rydberg series of a hydrogen atom (a), positronium (b), and a typical exciton (c).

levels of positronium as a function of the quantum number n. This set of energy levels is often referred to as a *Rydberg series*. The continuum at the top of the figure is the region of positive energies where the electron and the hole are so far away from each other that the Coulomb interaction no longer has an appreciable effect, and the energy is all of the kinetic type, $1/2(mv^2) = p^2/2m$, or energy of motion, where v is the velocity and $p = mv$ is the momentum.

The analog of positronium in a solid such as a semiconductor is the bound state of an electron–hole pair called an *exciton*. For a semiconductor the electron is in the conduction band and the hole is in the valence band. The electron and hole both have effective masses m_e and m_h, respectively, which are less than that m_0 of a free electron, so the reduced mass m_R is given by $m_R = m_e m_h/(m_e + m_h)$. When the electron effective mass is appreciably less than the hole effective mass, $m_e \ll m_h$, the relationship between them is conveniently written in the form

$$m_R = \frac{m_e}{1 + (m_e/m_h)} \tag{1.20}$$

which shows that for this case m_R becomes comparable to the electron mass. For example, if $m_e/m_h = 0.2$, then $m^* = 0.83\, m_e$. A comparison of the data in Table B.8 shows that this situation is typical for GaAs-type semiconductors. We also see from Table B.12 that the relative dielectric constant $\varepsilon/\varepsilon_0$ has the range of values $5.6 < \varepsilon/\varepsilon_0 < 15.7$ for these materials, where ε_0 is the dielectric constant of free space. Both of these factors have the effect of decreasing the exciton energy E_{ex} from that of positronium, and as a result this energy is given by

$$E_{ex} = \frac{m^*/m_0}{(\varepsilon/\varepsilon_0)^2} \frac{e^2}{4\pi\varepsilon_0 a_0 n^2} = \frac{13.6\, m^*/m_0}{(\varepsilon/\varepsilon_0)^2 n^2} \quad \text{eV} \tag{1.21}$$

as shown plotted in Fig. 1.32. These same two factors also increase the effective Bohr radius of the electron orbit, and it becomes

$$a_{eff} = \frac{\varepsilon/\varepsilon_0}{m^*/m_0} a_0 = \frac{0.0529\varepsilon/\varepsilon_0}{m^*/m_0} \quad \text{nm} \tag{1.22}$$

Using the GaAs electron effective mass and the heavy-hole effective mass values from Table B.8, Eq. (1.22) gives $m^*/m_0 = 0.058$. Utilizing the dielectric constant value from Table B.11, we obtain, using Eqs. (1.21 and (1.22), for GaAs

$$E_0 = 4.5\,\text{meV}, \quad a_{eff} = 12.0\,\text{nm} \tag{1.23}$$

where E_0 is the ground-state ($n = 1$) energy. This demonstrates that an exciton extends over quite a few atoms of the lattice, and its radius in GaAs is comparable with the dimensions of a typical nanostructure. An exciton has the properties of a

particle; it is mobile and able to move around the lattice. It also exhibits characteristic optical spectra. Figure 1.32 plots the energy levels for an exciton with the ground-state energy $E_0 = 18$ meV.

Technically speaking, the exciton that we have just discussed is a weakly bound electron–hole pair called a *Mott–Wannier exciton*. A strongly or tightly bound exciton, called a *Frenkel exciton*, is similar to a longlived excited state of an atom or molecule. It is also mobile, and can move around the lattice by the transfer of the excitation or excited-state charge between adjacent atoms or molecules. Almost all the excitons encountered in semiconductors and in nanostructures are of the Mott–Wannier type, except in organic nanocrystals, where they are Frenkel excitons.

PROBLEMS

1.1. Find the volume and the surface area of the cuboctahedron sketched in Fig. 1.6.

1.2. Consider a nanofilm formed from the two-dimensional close-packed lattice of Fig. 1.4. The smallest $n = 1$ nanofilm is, of course, one atom, and the $n = 2$ one is the seven-atom ($N = 7$) structure sketched on the left side of Fig. C.1 of Appendix C. The number of atoms around the edge is given by $N_{\text{edge}} = 6(n - 1)$ valid for $n > 1$, in analogy with Eq. (1.4). Make accurate sketches of the $n = 3$, $n = 4$, and $n = 5$ planar nanoparticles, and prepare the analog of Table 1.3 for this two-dimensional case for n values ranging from 1 to 5.

1.3. Evaluate the coefficients A, B, and C of the equation $N = An^2 + Bn + C$ for the number of atoms in a two-dimensional nanoparticle.

1.4. Find the distances between [111] planes of the zinc blende lattice in terms of the cubic unit cell dimension a for the following cases: (a) two nearest sulfur planes, (b) two nearest zinc planes, and (c) adjacent S and Zn planes. Why are there two answers for case (c)?

1.5. Consider three fcc nanoparticles containing approximately 1000 atoms: one in the shape of an octahedron, one in the shape of a cube, and one in the shape of the structure depicted in Fig. 1.9. Which has the largest, and which has the smallest value of surface energy per unit volume? What are their values?

1.6. A two-dimensional close packed lattice of circular atoms is illustrated in Fig. 1.4a. What is the radius of the interstitial sites of this lattice, and how numerous are they compared to the atoms?

1.7. A square of side a has zero-dimensional apices, four one-dimensional edges, and one two-dimensional face. If the square is projected out into the third dimension and the vertices of the two squares are connected with straight lines of length a, the resulting three-dimensional cube has 8 vertices, 12 edges, and 6 faces. Consider projecting the cube out into a fourth dimension and connecting the respective vertices with straight lines of length a, to form a four-dimensional "cube" called a *tesseract*. How many apices, edges, faces, and cubes does the tesseract have?

REFERENCES

1. N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, Saunders College, Philadelphia, 1976.
2. G. Burns, *Solid State Physics*, Academic Press, San Diego, 1985.
3. C. Kittel, *Introduction to Solid State Physics*, 7th ed., Wiley, New York, 1996.
4. P. Y. Yu and M. Cardona, *Fundamentals of Semiconductors*, 3rd ed., Springer-Verlag, Berlin, 2001.
5. T. P. Martin, T. Bergmann, H. Gohlich, and T. Lange, *Chem. Phys. Lett.* **172**, 209 (1990).
6. S. Sugano and H. Koizumi, *Microcluster Physics*, Springer, Berlin, 1998, p. 90.
7. C. P. Poole, Jr. and H. A. Farach, *Chemical Bonding in Semiconductor Physics*, K. Boer, ed., Wiley, New York, 2001, Vol. 1, Chapter 2.