Chapter 5
The Language of Electrons, Atoms, and Quanta

5.1 Classical Lorentz Theory

When we speak of oscillations at optical frequencies and their amplification, we are indeed a long way from the world of swinging pendulums and oscillating balance wheels. It is true that classical theory based on Newton's laws of motion and Maxwell's theory of electromagnetic radiation are inappropriate to deal with the interaction of radiation with atoms and molecules; for this we need the *quantum theory*. However, from a background of classical theory, certain aspects can be sketched in a semiclassical way, in which quantum ideas are superimposed on a classical base. Historically, this characterized the early development of the theory of radiation and the general features of the theory of *optical dispersion*. In this context "dispersion" refers to the dependence of the refractivity of a medium on the wavelength, which leads to the dispersion of, for example, white light by a glass prism into the colors of the rainbow.

Prior to the advent of quantum theory early in the last century, the interaction of radiation with matter was explained on the basis of the *electron theory* of H.A. Lorentz, in which the response of matter to an electromagnetic wave was expressed in terms of "atomic oscillators" pictured as electrons elastically bound to the atomic centers. The interaction of atoms with an electromagnetic wave was imagined as consisting in these electrons being driven into forced oscillation by the oscillating electric field component of the wave. It can be shown, however, that in order to have continuous absorption of energy from the wave (as opposed to a fleeting absorption when the wave first interacts with an electron, setting it in motion), it is necessary to assume that during the interaction with the wave, the driven electron oscillation must in effect experience a resistive force. This clearly cannot be a frictional force in the usual sense; and the *radiation reaction* force, which accounts for the energy radiated by the vibrating electron, proves to be too small to account for the degree of light absorption that can occur. Lorentz attributed the net absorption to the repeated interruption of the electron oscillation by collisions with other

F.G. Major, The Quantum Beat

atoms, resulting in the randomization of the oscillation phase. In the absence of this, the phase of the periodic electron velocity remains in quadrature with the driving force, and the average work done on the electron averages to zero over a period of oscillation of the field. The result would be that no net absorption of energy takes place. Through those phase-randomizing collisions, there is a continuous transfer of energy to the electrons that appears, through the same collisions, as random kinetic energy of the colliding atoms, that is, heat.

The same model was also used to describe the process of emission of radiation by atomic oscillators, when set into vibration by collisions with other atoms in an electrical discharge, or in a state of thermal agitation, as in a flame. It is well established classically, on the basis of Maxwell's theory, that an oscillating electric charge will radiate electromagnetic waves. In this case, since we have a negative charge (the electron) oscillating with respect to an equal positive charge, the wave that is generated is that of an oscillating electric dipole. This has a characteristic radiation pattern, that is, distribution of intensity in different directions, similar to that from a simple radio transmitter antenna. The frequency of the radiated electromagnetic wave is classically the same as the frequency of oscillation of the supposed atomic oscillator. If through some nonlinearity the atomic oscillator excitation results in some second or higher harmonics, at twice or a higher multiple of the fundamental frequency, the radiation will also contain those harmonic frequencies. It was one of the fatal flaws of classical theory in explaining atomic spectra that the observed frequencies emitted by atoms do not bear a simple harmonic relationship to each other.

5.2 Spectrum of Blackbody Radiation

But the breakdown in the classical theory of radiation, which finally led Planck to postulate the quantum of energy, first came in the explanation of the spectrum of the radiation in thermal equilibrium with matter, the so-called *blackbody radiation.*

This is radiation whose spectrum is characteristic of the equilibrium temperature, and it is independent of the nature of the matter interacting with it. It can be observed only under conditions where the interacting matter can thoroughly absorb and re-emit radiation at all frequencies. In practice this is achieved by studying the radiation inside an enclosure, which is provided with a small hole to allow a sample of the radiation to be analyzed outside the cavity. The observed continuous spectrum, showing the radiated intensity in a small fixed frequency band as a function of the center frequency of that band, is shown in Figure 5.1. Contrary to classical predictions, the graph tends to zero at the upper and lower ends of the frequency scale, with a maximum intensity at some intermediate frequency, which, in accordance with common experience, depends on the temperature: the color varies from red toward the blue as the temperature is raised. This is given precise expression in *Wien's displacement law*: the frequency at which the intensity is maximum shifts to higher values, in direct proportion to an increase in the temperature. Wien derived

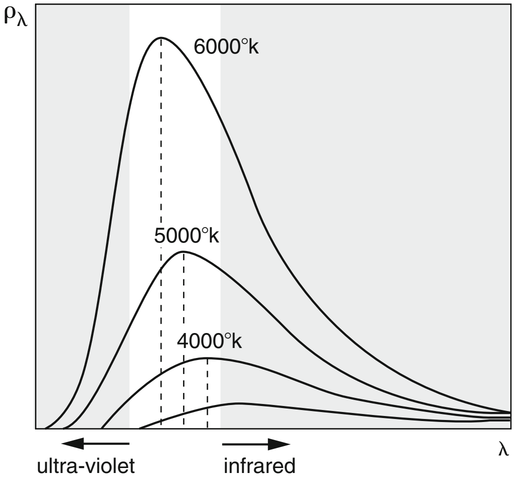

Figure 5.1 The spectrum of blackbody radiation showing the shift in the maximum with temperature

his law on the basis of classical arguments, prior to Planck's work, and it is borne out by experiment. The spectrum of sunlight is approximately that of a "black body" at a temperature of about 6000°K, with the maximum intensity occurring at around a wavelength of 0.5 µm, in the middle of the visible region of the electromagnetic spectrum.

5.3 The Quantum of Radiation: The Photon

After all attempts based on the classical theory of thermal equilibrium and the exchange of energy between radiation and matter failed to explain the observed spectrum, Max Planck in 1901 published a radically new theory, which was able to predict a spectrum in close agreement with experiment. It was based on the postulate that matter contained an immense number of electromagnetic "resonators" that could exchange energy with the radiation field not continuously in arbitrarily small amounts, but only in discrete units he called *quanta*, whose energy is proportional to the frequency: $E = h\nu$, where h is a universal constant of nature, now called Planck's constant, with a numerical value in our system of units of 6.6×10^{-34} joule · second.

A greater understanding of the physical processes that result in the emission of blackbody radiation came with the reinterpretation of the process by Einstein, who introduced the concept of a quantum of electromagnetic radiation, called a *photon*, which in some circumstances manifests a discrete particle nature. On this basis, blackbody radiation results when equilibrium has been reached between the *photon gas* and the atoms of matter through continual absorption and re-emission of photons by the atoms. When this model was applied to derive Planck's formula, it was found that the well-known *spontaneous* emission process, in which an atom gives up its energy of excitation spontaneously by emitting a photon, alone would not lead to an equilibrium consistent with Planck's formula. Einstein found it necessary to postulate that an atom that has absorbed a photon may not only re-emit it spontaneously, but may also be *stimulated* to re-emit it, with a probability that depends on the number of photons already present. When a group of such atoms or molecules undergo spontaneous emission, they do so independently of each other; there is, therefore, *no* correlation between the phases of their several contributions to the radiation emitted. In contrast, emission induced by existing photons, that is, *stimulated emission*, has a phase dictated by the phase of the existing radiation field, and hence all atoms subjected to this field will have more or less a common phase. This results in the radiation field remaining coherent in phase and increasing or decreasing in amplitude depending on whether the rate of emission is greater or less than the rate of absorption.

5.4 Bohr's Theory of the Hydrogen Atom

The success of the radically new quantum postulate of Planck soon saw the spread of quantum ideas to the hitherto intractable problem of explaining optical emission spectra of atoms. A wealth of accurate experimental data had been accumulated on the wavelengths of the many series of lines that form atomic spectra. Each chemical element emits its own characteristic, and for all but the simplest atoms, complex line spectrum. Intensive efforts had been devoted to finding regularities in these spectra, and a number of empirical rules were enunciated, all of which brought some order to their practical analysis.

The theoretical breakthrough came after the success of the *nuclear model* of the atom, which was postulated by Ernest Rutherford around 1911 to explain unexpectedly large angles of scattering of high-speed α-particles (a product of natural radioactivity of certain elements) by atoms in a gold foil target. The model was strikingly confirmed in subsequent years in his laboratory, a feat he announced as possibly more important than the outcome of what was then called the Great War. For this he received the Nobel Prize and given the title Lord Rutherford of Nelson (his birthplace in New Zealand). Prior to that, there was intense speculation as to just how electrons and protons, the elementary particles known at the time, were arranged in atoms. As is now familiar to everyone, the basic arrangement is that almost the entire mass of an atom resides in a small, positively charged nucleus, which is surrounded by a cloud of negatively charged electrons.

Around 1913, the Danish physicist Niels Bohr, by a set of ad hoc quantum notions superimposed onto a classical planetary model of the hydrogen atom, was able to obtain with remarkable accuracy the wavelengths of a series of lines in the spectrum of that atom. The most radical of Bohr's postulates was that there exist certain orbits in which electrons could circulate indefinitely without radiating energy, contrary to the classical prediction that an orbiting charge should lose energy by radiation and eventually spiral into the nucleus. These he called *stationary* orbits, and he postulated that of all the possible orbits that classical theory allows, only those are stationary that satisfy the following condition on their angular momentum L:

$$L = \frac{nh}{2\pi}, \tag{5.1}$$

where n is an integer and h is the same constant Planck had used to define the quantum of energy. For a circular orbit of radius r, $L = mVr$, where m is the mass of the electron and V its velocity. He further postulated that the frequency of radiation emitted by the atoms is not the vibration or rotational frequency of the electron in the classical sense, but is derived from Planck's formula. Thus, when an atom makes a quantum transition from a stationary state of energy E_2 to one having energy E_1, the frequency of the radiation is that of the radiated quantum, that is,

$$\nu = \frac{(E_2 - E_1)}{h}. \tag{5.2}$$

Of course, these radical postulates were not made lightly. The line spectra of atoms show remarkable regularities, with series of lines forming striking patterns, plausibly reminiscent of the classical vibration spectra of complex structures. It would be natural to assume that these vibration spectra should form the basis of an explanation of the spectrum. Unfortunately, of all of the precise experimental data that was available and some empirical formulas that were discovered relating the wavelengths in the spectra, none was consistent with the harmonic relationships characteristic of classical vibration frequencies.

Bohr's ad hoc postulate identifying stationary orbits became a little less so through the work of de Broglie, published in 1924. In this de Broglie argued on the basis of the dual particle–wave nature of light, which was then the subject of much speculation and debate, that material particles have the same duality. The success of Bohr's theory seemed to hint at a wave property of electrons, since at the time the only context in which equations contained integers was in normal modes of vibration, and the interference of waves. On the basis of the special theory of relativity de Broglie was able to find the connection between the particle and wave nature of all matter and radiation, in a theory called *wave mechanics*, the precursor of quantum mechanics. According to de Broglie, a particle of mass m moving with a velocity V has a *wave* associated with it "guiding" its motion, whose wavelength, now called the *de Broglie wavelength*, is given by $\lambda = h/mV$, where h is, as usual, Planck's constant. If we use this result in Bohr's equation for the stationary orbits, we find $(h/\lambda)r = nh/2\pi$; that is, $2\pi r = n\lambda$. But this is precisely the condition for a resonant mode of vibration of a circular string supporting oscillations with a wavelength λ; any other radius would not have the wave reinforcing itself as it traveled around the circle.

The theory of Bohr, elaborated by Sommerfeld, and now referred to as the "old quantum theory," dealt only with "stationary" quantum states and quantum numbers; it had little to say about nonstationary phenomena such as transitions between states and collisions between particles. This situation changed with the coming of quantum mechanics.

5.5 The Schrödinger Wave Equation

The spirit of de Broglie's description remains in the subsequent quantum theory of Schrödinger. The concept of a wave determining the motion of a particle implies the radical notion that the amplitude of a wave, given as a function of the coordinates and called a *wave function*, is to be used to describe the motion of a particle, rather than regarding a particle as a point mass occupying a certain position in space specified by its coordinates. The physical interpretation of the wave function, conventionally represented by the Greek letter ψ, lends itself to some speculation in the minds of some, hinting at a mysterious wave that guides the motion of matter. However, a probabilistic view prevails in which $|\psi(x, y, z)|^2$ is taken as the space density of the *probability* that the particle is at the coordinates x, y, z, in the sense that $|\psi(x, y, z)|^2 dx\ dy\ dz$ is the probability of the particle being found in a cell of sides dx, dy, and dz centered at the point (x, y, z). Since the particle must be somewhere with a 100% certainty, it follows that the wave function must satisfy the following *normalization* condition:

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} |\psi|^2 dxdydz = 1. \tag{5.3}$$

This, of course, imposes a mathematical restriction on the *wave function*: Its integral must be finite.

In Schrödinger's *wave mechanics*, which is one mathematical representation of quantum mechanics, the equations of motion of classical mechanics are replaced by a differential equation, called the *Schrödinger equation*, to determine the *wave function*. Thus, for example, the equation for a free electron having energy E in a one-dimensional world would be as follows:

$$\frac{d^2\psi}{dx^2} + \frac{8\pi^2 mE}{h^2}\psi = 0. \tag{5.4}$$

Of all the mathematical solutions of the Schrödinger equation, those that may be accepted as representing the stationary states of a physical system are defined as those particular solutions, called *eigenfunctions* (German for proper functions), that are finite and satisfy certain conditions at the boundaries of the system. For example, if an electron obeying the above equation is confined between two plane, parallel, "impenetrable walls" forming the boundaries at $x = 0$ and $x = L$, the stationary solution of the Schrödinger equation describing that electron would be equal to zero at those points and beyond. One can readily verify that the following are solutions:

$$\psi_n = N \sin(k_n x), \tag{5.5}$$

where

$$K_n = n\frac{\pi}{L}, \quad E_n = n^2 \frac{h^2}{8L^2 m}, \tag{5.6}$$

and $n = 1, 2, 3, \ldots$. We note that were it not for the boundary conditions $\psi(0) = 0$ and $\psi(L) = 0$, the equation would have been satisfied by $\sin(kx)$, where k, and therefore E, are continuous variables, and not "quantized" to have the discrete values labeled with the index n: k_n and E_n. The functions $N \sin(k_n x)$ are the stationary wave functions, the eigenfunctions of Schrödinger's equation for the particular system we have assumed. They are analogous to the classical normal modes of vibration of a system.

For the 3-dimensional case of a particle confined in a rectangular box with sides L_1, L_2, L_3, the eigenfunctions have the form

$$\psi_{l,m,n} = \sqrt{\frac{8}{L_1 L_2 L_3}} \sin(k_l x) \sin(k_m y) \sin(k_n z), \tag{5.7}$$

where

$$k_l = \frac{l\pi}{L_1}, k_m = \frac{m\pi}{L_2}, k_n = \frac{n\pi}{L_3}, \tag{5.8}$$

and the quantum energy levels are given by

$$E_{l,m,n} = \frac{h^2}{8\pi^2 m}\left[k_l^2 + k_m^2 + k_n^2 +\right]. \tag{5.9}$$

We note that we now have three quantum numbers l, m, and n to distinguish the various possible stationary states, and that these appear in the quantization of the components of the wave vector **k** along the three coordinate axes. If we recall the formula for the de Broglie wavelength, we find that $k = (2\pi/h)mV$; that is, it is the linear momentum that is quantized. The constant factor $\sqrt{8(L_1 L_2 L_3)}$ is introduced to meet the normalization condition.

We note that the stationary states we found for an electron in a box are far from the classical picture of a point mass bouncing back and forth between the boundaries. A particle moving back and forth would be represented as a time-dependent wave function that, at any moment is small everywhere except in the neighborhood of the particle position. Such a wave function, called a *wave packet*, can be synthesized as a sum over the harmonic eigenfunctions, following the spirit of the Fourier expansion theorem. Each eigenfunction corresponds to a different energy, and therefore a different frequency (since Planck's formula $E = h\nu$ still holds), with the result that the wave packet will have a time dependence reflecting the motion of the particle.

5.6 Quantum Numbers of Atomic States

If a particle is subjected to a central force, that is, one directed toward a fixed point, such as the electrostatic Coulomb force that a nucleus exerts on the electrons surrounding it in an atom, three quantum numbers will again be required to specify a stationary state. In this case the spherical symmetry of the equation suggests that the solution is most naturally expressed using the spherical polar coordinates r, θ, ϕ. The quantum numbers conventionally designated as n, l, m play a role in close analogy to the indices used to label the various normal modes of vibrations of a sphere. The values of the quantum numbers are restricted as follows: $n = 1, 2, 3, \ldots$, while $l \leq (n - 1)$ and $m = l$, $(l - 1)$, $(l - 2) \ldots -(l - 2)$, $-(l - 1)$, $-l$. The part of the wave function that is a function of the r-coordinate has a number of nodes (zeros) given by $(n - l - 1)$, and the part that is a function of the co-latitude θ has $(l - m)$ nodes off-axis. Following spectroscopic convention, electrons in an atom having $l = 0, 1, 2, 3, \ldots$ are called s-, p-, d-, f-electrons, etc., respectively. The quantum numbers l and m, which are associated with the angular part of the wave function, in fact reflect the quantization of the angular momentum and its component along the polar axis, respectively. In Figure 5.2 is shown the probability distribution for a particle in a central field in the $n = 3$, $l = 2, m = 0$ quantum state. According to the theory, a system having nominally an orbital momentum quantum number l actually will have orbital angular momentum of $\sqrt{l(l + 1)}$ in units of $h/2\pi$, whereas the maximum component along the polar axis is only l. (We will usually omit the unit $h/2\pi$ unless we are doing a numerical calculation.) Thus the theory predicts that the maximum component the angular momentum can have along any given axis is somewhat less than the magnitude

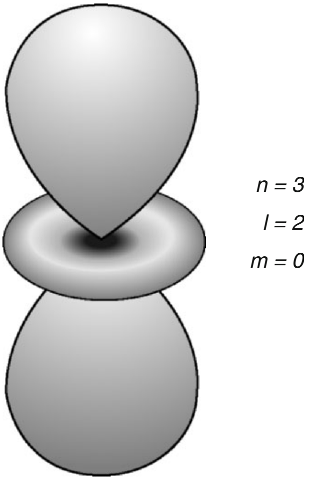

Figure 5.2 Example of the probability distribution for a particle in a central field

of the angular momentum itself. This is a strictly quantum effect, since classically the angular momentum is a vector that can assume any direction, and in particular can point exactly in the direction of any given axis. The effect can be interpreted in terms of vectors by saying it arises from a quantum uncertainty in the angle the angular momentum vector makes with the axis. The quantization of the component of angular momentum along an axis arises mathematically from the condition that the solution to the Schrödinger equation must be a simple periodic function of the angle ϕ around the axis; that is, it must repeat itself every 360°. This physical requirement imposed on the mathematical solution bears some resemblance to the implied condition for normal modes in the Bohr circular orbits. In any event, it constitutes one of the most radical breaks with classical mechanics: It implies that the angular momentum of a system can only assume certain discrete orientations with respect to a given axis; this is sometimes called *space quantization*, and it is of profound importance in the quantum theory of atoms subjected to an external magnetic field, and the attendant shifts in energy levels: the *Zeeman effect*.

Since an atomic angular momentum will have associated with it a magnetic moment (both due to the orbital motion of the charged electron and its intrinsic spin), the energy shift produced by a magnetic field is expected to depend on the component of the angular momentum along the field and therefore, with the axis chosen along the field direction, on the quantum number m. For this reason m is called the *magnetic* quantum number, and to reiterate, for a state with a *total* angular momentum (including spin) quantum number J (which, as we shall see, may be integral or half-integral) the magnetic quantum number m_J, can have one of the following $(2J + 1)$ discrete values: J, $(J - 1), (J - 2), \ldots, -(J - 2)$, $-(J - 1), -J$. For example, a particle in an angular momentum state described nominally as a $J = 5/2$ state may have as its component along a given axis one of the following values: $+5/2, +3/2, +1/2, -1/2, -3/2, -5/2$.

5.7 The Vector Model

It should be emphasized that the quantum numbers, while they represent in quantum mechanics the results of measurement of a particular dynamical quantity, such as angular momentum, it is only in systems involving very large quantum numbers that they approximate classical behavior. It happens that we can, according to what is called the *vector model*, retain the concept of angular momentum as a classical vector, provided that we give these vectors properties that are peculiar to quantum mechanics. The uncertainty in pointing the angular momentum exactly along a given direction is one of them. The other concerns combining different angular momenta to obtain a resultant: the result of adding two angular momenta whose quantum numbers are, for example J_1 and J_2, where $J_2 < J_1$, would classically be *any* value between $(J_1 - J_2)$ and $(J_1 + J_2)$, depending on the angle between the two angular momentum vectors, whereas in quantum theory the resultant is one of a discrete set that starts with $(J_1 - J_2)$ and by increments of one

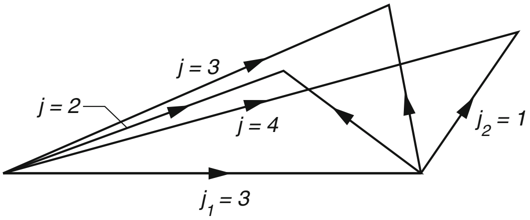

Figure 5.3 Vector diagram for the addition of angular momenta

unit reaches $(J_1 + J_2)$. For example, suppose a system in an angular momentum state with quantum number $J_1 = 3$ and another with $J_2 = 1$ interact in such a way that results in stationary states of the combined angular momenta. The quantum numbers belonging to this combined representation would be 2, 3, 4, implying according to the vector model that the angular momentum vectors can make only certain discrete angles with respect to each other, as illustrated in Figure 5.3.

If the particle is an electron, a complete specification of its quantum state requires not only the dependence of its wave function on the space coordinates, but also the state of another attribute of the electron called the *spin*. This is an *intrinsic* angular momentum of $^1/_2(h/2\pi)$, part of what it is to be an electron. It was first introduced to explain atomic spectra and later brilliantly shown by Dirac to be a logical necessity, forming an integral part of a relativistic quantum theory. For a free electron, the spin component along any given axis can only be $+1/2$ or $-1/2$, corresponding to only two possible directions of spin.

For electrons in an atom, a total angular momentum larger than 1/2 can result from the spin combining with the orbital angular momentum of its motion around the nucleus, which is conventionally represented by l and is always integral. It can be shown that the magnetic field produced by the orbital motion of the electron can exert a torque on its own spin, a coupling called the *spin–orbit interaction*, which is extremely important in understanding atoms. In the absence of other torques acting separately on the two types of angular momentum, such as a strong external magnetic field, the two will give a resultant angular momentum represented conventionally by j, which is conserved, and quantized both in magnitude and spatial orientation. For example, an electron in an orbit with orbital angular momentum $l = 2$ will have a resultant, when combined with its spin of 1/2, equal to either $2 + 1/2$ or $2 - 1/2$, that is, 5/2 or 3/2. Recall that these numbers give the maximum component observable along any given axis in units of $h/2\pi$.

5.8 The Shell Structure of Electron States

When there is a large number of electrons in different orbits, the prediction of the possible combined angular momentum states quickly becomes very complicated; not only is there spin–orbit coupling, but also interactions between the spin and orbital magnetic moments of different electrons. Fortunately, it happens that electrons in an atom can be grouped into *shells*, which, as we shall see, can contain only a certain maximum number of electrons. When completely filled, a shell has zero resultant angular momentum; so that only electrons in any incomplete shells need be considered in arriving at the overall atomic angular momentum state.

The reason for venturing a little into the abstruse realm of quantum theory is that it is essential for any basic understanding of atomic and molecular structure and dynamics. We recall that the atoms of the chemical elements have small positive nuclei, where most of the mass resides, surrounded by a cloud of negative electrons that occupy available quantum *states*, each state labeled by a set of three quantum numbers, plus a fourth specifying the spin state. We have already seen that for a given value of the quantum number l there are $(2l+1)$ states with different m; if we include the two possible directions of the spin, this number is doubled. (The presence of spin–orbit coupling requiring a description in terms of the total (spin plus orbital) quantum numbers does not affect the number $2(2l + 1)$.) These states correspond to different orientations of the orbital and spin angular momenta with respect to a fixed axis. In the absence of an external field, such as a magnetic field, all directions in space are identical, and the energy of electrons in these states is the same; they are all at one energy level. They are called *degenerate* states. Furthermore, it is found that for a pure Coulomb (inverse square law) electrostatic field, such as we have in the hydrogen atom, the solution to the Schrödinger equation yields possible values of energy that depend only on the quantum number n, and so there is degeneracy with respect to the l quantum number as well. Now, for each value of n, the quantum number l can assume any of $(n - 1)$ values, and as we have seen, to each l value there are $2(2l + 1)$ degenerate states. The total number of degenerate states having the same n is therefore

$$\sum_{0}^{n-1} 2(2l+1) = 4\frac{n(n-1)}{2} + 2n = 2n^2. \qquad 5.10$$

The common energy of these states can be shown to be

$$E_n = -\left(\frac{2\pi^2 m Z^2 e^4}{h^2}\right)\left(\frac{1}{n^2}\right), \qquad 5.11$$

in agreement with the old quantum theory of Bohr, which was already known to be in remarkable agreement with experiment. There is an infinite number of energy levels corresponding to n ranging from 1 to ∞; Figure 5.4a shows some of the lower states. Unlike the Bohr theory, the electrons are not localized along particular orbits but must be regarded as spread out with a radial density given by $4\pi r^2|\psi|^2$, which is illustrated for several states in Figure 5.4b. We note that the average radius increases with n, the outermost electrons having the highest n and the smallest binding energy.

For atoms having a large nuclear charge, and therefore many electrons, the exact solution of the Schrödinger equation becomes impossible, and approximate

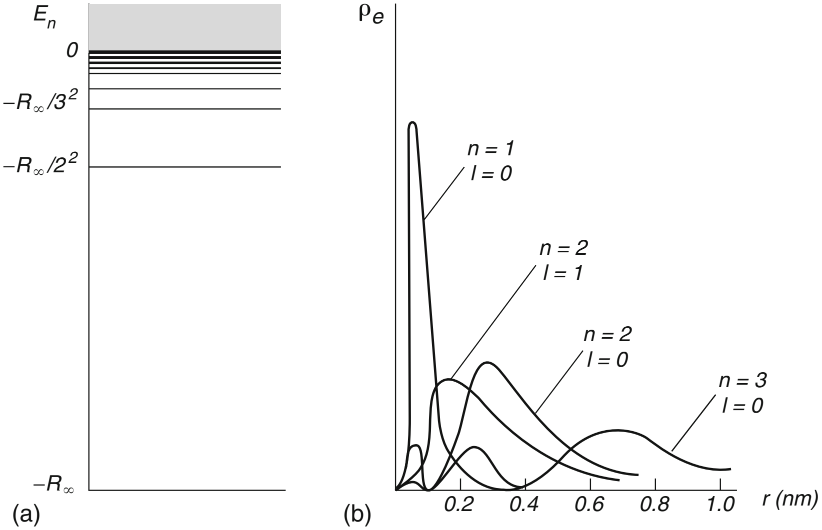

Figure 5.4 (a) The energy levels of the H-atom, and (b) the radial dependence of some of the lower energy wave functions

numerical methods have been developed. An approximation that has proved very useful is to assume that each individual electron moves in an electrostatic field produced by the nuclear charge and an average *spherically symmetric* distribution of charge due to the other electrons. Of course, after solving the Schrödinger equation using this approximate field and obtaining the charge distribution of each electron from its calculated wave function, the combined charge distribution so derived must agree with the one assumed in the first place. The important point for us is that if the field acting on the electrons can indeed be taken to be spherically symmetric, then the same quantum numbers n, l, m can still be used with the same significance, except that the radial distribution of electrons is no longer purely hydrogen-like, and the energy is no longer a function of the principal quantum number n alone, but depends on l as well. That is, the l-degeneracy is removed. However, the dependence of the energy on l is still generally weaker than that on n, aside from some important exceptions for larger l-values. The m-degeneracy remains, and levels are grouped around the different l-values; these groupings are the *shells* mentioned earlier. It is one of the early triumphs of quantum theory that it was able to predict the number of quantum states in each shell. Thus, for example, states having $n = 4$ and $l = 2$ would be said to belong to the $4d$ shell, and those having $n = 5, l = 0$ are in the $5s$ shell; the former would number $2(2 \times 2 + 1) = 10$ states, and the latter just $2(2 \times 0 + 1) = 2$ states.

5.9 The Pauli Exclusion Principle

Starting with a nucleus having a given number of protons, in order to construct a neutral atom in its ground state we must take the same number of electrons and allocate them one by one to progressively higher-energy quantum states beginning with the lowest-energy state first. This atomic building principle is based on the condition that no two electrons can occupy the same quantum state, that is, have the same set of quantum numbers. This is a statement of the *Pauli exclusion principle*, which is at the heart of the quantum explanation of atomic structure and spectra. It can be deduced from a symmetry property of *wave functions* representing a system of electrons and some other elementary particles. Since individual electrons are indistinguishable, in the sense that we cannot know which electron occupies a particular position and spin state, an exchange of the assignment of these between any two electrons in the wave function ψ cannot change the observable $|\psi|^2$. Therefore, an electron exchange must either leave ψ unchanged (symmetric wave function) or at most change its sign (antisymmetric wave function). It happens that photons have the former symmetry, while electrons the latter. For electrons this means that the probability of finding two electrons in identical states is zero, since in that event an exchange of the two electrons must on the one hand leave the wave function unchanged, but on the other its sign must change; this can happen only if it is zero. Once an electron occupies a certain state, that state is said to be filled. This means that in constructing the ground state of an atom, each state must be filled before the next higher energy state is filled. The assignment of electrons to the different possible quantum states is analogous to the assignment of passengers to *single-*occupancy berths on a cruise ship; each berth has a number, and the fare schedule is based mainly on which deck the berth is located, with some differences within a given deck depending on its location. For the electrons in an atom, the "decks" are the shells, and the "fare" is the energy. Unlike a cruise ship, however, the electrons of an atom are in the stable ground state when their total energy ("fare") is a *minimum*.

Since we shall be concerned with crystalline quartz (SiO_2) in the next chapter, let us consider the elements oxygen and silicon as examples. They have (positive) nuclear charges of 8 and 14 respectively, in units of electronic charge. Therefore, oxygen will have the shells 1s, 2s filled and be two short of filling 2p, while silicon will have the 1s, 2s, 2p filled and have two electrons in each of the 3s and 3p outer shells.

It is the outermost electrons in an atom that determine its chemical properties and its interaction with radiation in the optical region of the spectrum. The inner electrons are unable to take part in any small exchange of energy, since all neighboring energy states are filled. Of course, if sufficient energy is involved, as in electron bombardment in an X-ray tube, inner electrons do play a part; but ordinary chemical reactions and optical transitions involve relatively little energy. The Mendeleev periodic system of the chemical elements finds a ready explanation in terms of the filling of shells as the nuclear charge (atomic number, Z) is incremented. Thus the property of having a completely filled outer p-shell corresponds to the noble gases and will recur at $Z = 2$ (He), $Z = 2 + 8$ (Ne), $Z = 2 + 8 + 8$ (Ar), etc. Next would be the alkali elements with a single electron outside a closed shell; they are at $Z = 1$ (H), $Z = 2 + 1$ (Li), $Z = 2 + 8 + 1$ (Na), $Z = 2+8+8+1$ (K), etc. Then the alkaline earths, Be, Mg, Ca, Sr, ..., with two electrons outside closed shells, and so on. This simple progression is interrupted when we reach a point where it becomes "cheaper" in energy to go to a higher n-value than to add to a shell with a high l-value. This leads to the so-called *transition elements*, for example, those involved in filling the 3d shell (*after* the 4s shell has been filled), Mn, Fe, Co, Ni.

In all the elements, the inner closed shells and the nucleus form a tightly held inner core, with an unbalanced positive charge equal to the charge of the outer electrons. In the context of chemical bonding, the outer electrons are referred to as the *valence* electrons, of which silicon has four and is therefore tetravalent, and of which oxygen lacks two to complete a shell and is thus divalent. Without going into the subject any more deeply than we absolutely have to, we will simply state that the bonding between atoms to form compounds can be characterized according to the extent that the valence electrons (a) overlap between the atoms (covalent character) or (b) are transposed from one atom to the other, forming positive and negative ions that attract each other (ionic character). Whether the bond between a particular pair of atoms is predominantly covalent or predominantly ionic depends on the relative energy "cost" of the electrons arranging themselves according to the one or the other; recall that stability belongs to the lowest energy. The covalent bond may involve one valence electron, as in the bond between Si and O in quartz, or more than one electron, as typified by the bond between C and O in carbon dioxide (CO_2), in which the carbon atom has a double bond with each oxygen atom.

Now, in a covalent bond, where the dominant feature is the overlap of valence electrons belonging to the two atoms (recall that the electrons are to be viewed as smeared over all space according to the magnitude of their wave function), it is reasonable to expect that the possible distribution of the valence electrons around the inner core will determine the directions along which the bonds occur.

5.10 Spectroscopic Notation

A central problem in the quantum mechanical treatment of atomic observables is to find how the angular momenta of the constituent particles must be coupled in order that the energy and angular momentum are simultaneously in stationary quantum states. Because of the magnetic interactions between the particles, the individual particles will not maintain a constant direction with respect to some fixed axis and cannot define a "stationary" quantum state. In a system comprised of many interacting particles, the total angular momentum of the system will always be conserved, remaining constant in magnitude and direction, like an ideal gyroscope.

It may also happen that the angular momenta of particles within subsets of the total may be coupled to form conserved parts of the total angular momentum. The magnitude of such conserved angular momenta and their components along an arbitrary axis can serve to describe a stationary quantum state. Thus in one scheme of coupling angular momenta of electrons in a complex atom, called the *Russell–Saunders* coupling, the orbital angular momenta of the electrons are combined, then separately all the spin angular momenta are combined, and finally a resultant of the total orbital and spin angular momenta is obtained.

We recall that in combining angular momentum in quantum theory, we may use the vector model representation, provided that we remember that we are dealing with quantum numbers and that special quantization rules must be observed. Let us consider two examples that will be of considerable interest to us later: the alkali atoms rubidium and cesium. In their ground state, they have only one electron outside closed shells. In the ground state we are considering, this electron has no orbital angular momentum and therefore only the spin angular momentum of 1/2, with two possible components along a given axis, $+1/2$ or $-1/2$, and $g = 2$. If this single outer electron occupies the next higher energy state, it would have an orbital angular momentum of one unit, that is, $l = 1$, in addition to its spin. These angular momenta are not individually constant in direction, but the total angular momentum is conserved; according to quantum rules, the total can be only $J = 1/2$ or $J = 3/2$. Because of the relative weakness of the magnetic interactions compared to electrostatic pull of the nucleus, there is a difference in energy between these two states much smaller than would accompany a change in orbit, and this difference is therefore called the *fine-structure* splitting. It is due in this case to the spin–orbit interaction we mentioned earlier in this chapter.

The notation used by spectroscopists to designate these two states in the alkali atoms is $^2\mathrm{P}_{1/2}$ and $^2\mathrm{P}_{3/2}$. The letter indicates that the orbital angular momentum $L = 1$, the superscript 2 is the value of $(2\mathrm{S} + 1)$, where S is the spin angular momentum (in this case $\mathrm{S} = 1/2$), and finally, the subscripts 1/2 and 3/2 are the two values of total angular momentum J. In this notation the ground state is designated as $^2\mathrm{S}_{1/2}$.

5.11 The Hyperfine Interaction

The electron is by no means the only fundamental particle with intrinsic spin and magnetic moment; both the proton and neutron, which are the constituents of atomic nuclei, also have these attributes. These particles have the same magnitude of spin as the electron, but since their charge-to-mass ratio is 2000 times smaller, we would expect, at least classically, that their magnetic moment is also smaller in approximately the same ratio. In fact, as with the magnetic moment of the electron, classical theory is inapplicable, but the classical moments are used as units; for the electron it is the *Bohr magneton*; here it is the *nuclear magneton*. As with the electron, the magnetic moments of the proton and neutron are expressed in terms of g-factors defined as follows: $\mu = g_n I \mu_n$, where μ_n is the classical value of the magnetic moment of a particle with the charge and mass of a proton and an angular momentum of one unit, $h/2\pi$. The measured value for the proton is about $g_p = 5.586$ and for a free neutron $g_n = -3.82$. Again we see that classical theory is invalid, particularly for the neutron, which, being neutral, should have no magnetic moment at all. The question of what spin and magnetic moment a particular nucleus as a whole exhibits is a complicated one of nuclear structure, involving in general a large number of interacting protons and neutrons. The existence of a nonzero nuclear spin, which like total electronic angular momentum is limited to integral or half-integral values, further complicates the question of the angular momentum states of an atom, since the nuclear magnetic moment associated with it can interact with that of the outer electrons. Since the magnetic moment of the nucleus is so much weaker than the electron moment, it is expected that the different possible orientations of the nucleus will lead only to narrow splitting of the energy states. In recognition of that fact, the interaction between the electron and nucleus is referred to as the *hyperfine* interaction. It is precisely transitions between states separated by this hyperfine interaction that give rise to the sharp resonances used in the atomic standards in the microwave region of the spectrum. The assignment of angular momentum quantum numbers to the quantum states of an atom is very much affected by the addition of the nuclear spin, with important consequences, as we shall see, for any process involving exchange of angular momentum between an atom and radiation, for example.

It is a remarkable fact that in the $^2S_{1/2}$ ground state of the alkali atoms, the electron has zero orbital angular momentum, which classically would be interpreted as a collapsed electron orbit passing right through the nucleus; even the quantum picture is one of an electron spread out in a spherically symmetric way around the nucleus, with a finite probability of being found in the nucleus itself. A thoughtful reaction to this revelation might be, Why is there no nuclear reaction between the electron and the particles that make up the nucleus. The answer is that electron capture by the nucleus can occur in some species of nuclei; but where it is allowed, it is far more likely to involve the innermost electrons in the atom, in a process called K-capture because the innermost shell of an atom is called the K shell. Unlike s-electrons, all others in $l = 1, 2, 3 \ldots$ orbital angular momentum states have a vanishingly small probability of being in the nucleus. The spherical symmetry of the electron distribution in the $^2S_{1/2}$ state and its finite value in the nucleus have an important bearing on the computation of the interaction energy between the magnetic moment of the nucleus and that of the electron. They mean that we are not dealing with two separated magnetic dipoles, like two little magnets interacting with each other; rather it is a magnetic dipole embedded in a magnetized, spherically symmetric medium, as shown in Figure 5.5.

The problem is to compute the amount of energy that would be required to remove the embedded magnet from the center of that magnetized medium. Classically, reversing the relative directions of the magnetization of the magnet and medium merely changes the sign of the energy, the interaction changing from one

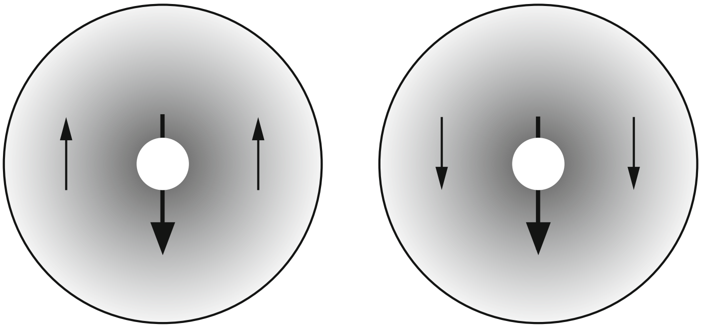

Figure 5.5 The magnetic moment of the nucleus interacts with that of the electron cloud surrounding it

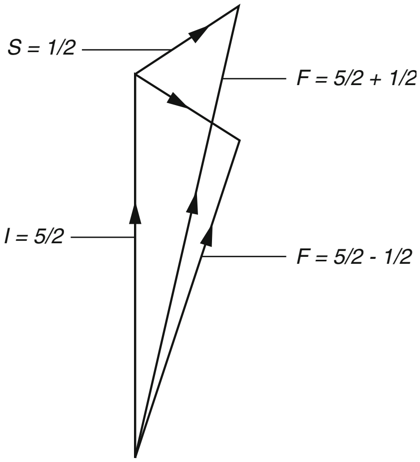

Figure 5.6 The quantum addition of angular momenta 5/2 and 1/2 according to the vector model

of attraction to one of repulsion; however, as we have become accustomed by now, this contradicts quantum mechanics. Simply put, the two possible angular momenta given nominally as $I + 1/2$ and $I - 1/2$ cannot be regarded as having the relative directions of the nuclear and electron spins reversed. Figure 5.6 illustrates the addition of an angular momentum of $5/2$ with one of $1/2$ according to the vector model. We see that since the magnitudes of the vectors have the form $\sqrt{5/2(5/2+1)}$ and $\sqrt{1/2(1/2+1)}$, the vectors for the angular momenta, which are nominally $(5/2 + 1/2)$ and $(5/2 - 1/2)$, do not have the $1/2$ angular momentum in opposite directions relative to the $5/2$.

The quantum-mechanical solution to the problem of the magnetic interaction between a nuclear moment and an overlapping electron distribution is associated with the name of Fermi, who obtained it as an early application of what was then the new quantum mechanics. The expression he obtained for the energy, in terms of the probability density of the electron at the nucleus and the magnetic moments of the nucleus and electron, is as follows:

$$E = \left(\frac{8\pi}{3I}\right)\mu_e\mu_n|\psi(0)|^2\left[F(F+1) - I(I+1) - J(J+1)\right], \qquad 5.12$$

where $|\psi(0)|^2$ represents the electron density at the nucleus. For zero orbital angular momentum states having the same total electron angular momentum J, we can write for the energy separation between adjacent F values the following:

$$E(F) - E(F-1) = \frac{16\pi}{3}\mu_e\mu_n|\psi(0)^2|\left(\frac{F}{I}\right). \qquad 5.13$$

The application of these formulas to such complex atoms as rubidium and cesium is not expected to yield very accurate results, since many simplifying assumptions have been made; among the more serious are these: A point magnetic dipole was assumed for the nucleus, as was a single electron in an unperturbed state. Even for the hydrogen atom, where these assumptions should be far more tolerable, the drive for accuracy in the theoretical ground state hyperfine separation has led to ever more sophisticated higher-order corrections being computed. As we shall see, thanks to the hydrogen maser this hyperfine separation in hydrogen is undoubtedly the most accurately measured quantity in physics: to better than twelve significant figures! One of the early triumphs in this field was the evidence that there was an "anomaly" in the magnetic moment of the electron; the value deduced experimentally did not agree with the then most advanced relativistic theory of the electron, the Dirac theory, which predicted that the electron g-factor should be exactly 2. In fact, it was found that $g = 2(1.00114\ldots)$, a number that has been the subject of precise studies by Dehmelt et al. (Dehmelt, 1981).

In the case of the rubidium atom, there are two naturally occurring *isotopes*, that is, atoms having the same electronic structure (which identifies them as rubidium) and therefore the same nuclear charge, but with a different nuclear mass because of a difference in the number of neutrons (see Figure 5.7). Natural rubidium is about 72% mass 85 with nuclear spin $I = 5/2$ and 28% mass 87, which has an extremely weak radioactivity and nuclear spin $I = 3/2$. If we follow the quantum rules for combining angular momentum, we will find that the ground state of Rb^{85} splits into energy levels with angular momenta equal to $(5/2 - 1/2)$ and $(5/2 + 1/2)$; that is, $F = 2$ and $F = 3$. Note that we can write symbolically $\mathbf{J} = \mathbf{L} + \mathbf{S}$ and $\mathbf{F} = \mathbf{J} + \mathbf{I}$ to represent the (vector) addition of orbital and spin angular momentum to obtain the total electronic angular momentum, and then the addition of the nuclear moment **I** to get the total conserved angular momentum **F**.

The assignment of angular momentum quantum numbers to the first energy level above the ground state in Rb^{85} is somewhat more complicated, since we have to combine $J = 3/2$ with $I = 5/2$ in addition to the combination of $I = 5/2$ with

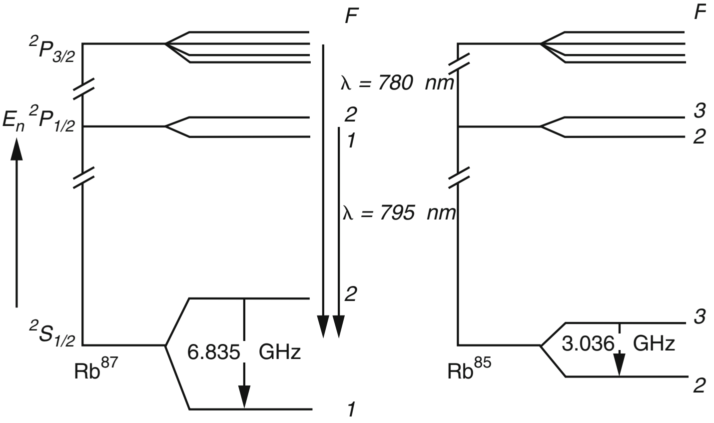

Figure 5.7 Hyperfine structure of low lying states in Rb^{85} and Rb^{87}

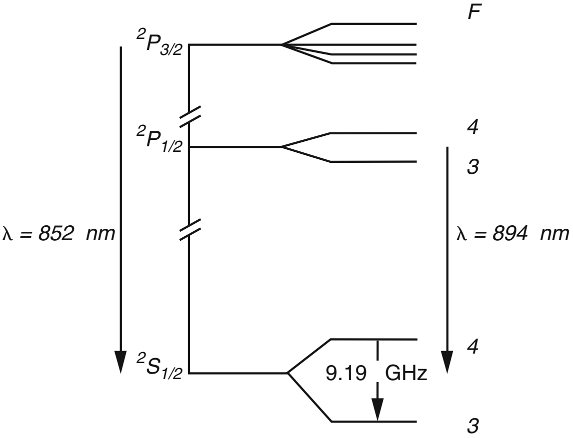

Figure 5.8 Hyperfine structure of low-lying states in Cs^{133}

$J = 1/2$, which leads to the values we have already found for the ground state. In general, we simply write all values between $I + J$ and $I - J$, that is, $F = 4, 3, 2, 1$.

Similar arguments may be used to find the angular momenta for the ground state and first excited states of the cesium atom (see Figure 5.8). There is only one stable isotope of cesium, mass 133, with a nuclear spin $I = 7/2$. Hence in the electronic ground state, which has $J = 1/2$, the possible total angular momenta are $F = 4$ and $F = 3$. For the first excited electronic state, which has two electronic angular momentum states, $\text{J} = 1/2$ and $\text{J} = 3/2$, the coupling with the nuclear spin leads to $F = 4, 3$ for the first J value and $F = 5, 4, 3, 2$ for the other J value.

As already indicated, the magnetic interaction of the nuclear moments with the electrons is expected to be very small compared to the other interactions that determine the quantum energy levels of an atom. Nevertheless, it is precisely the magnetic hyperfine separations in the ground states of rubidium and cesium that have come to be distinguished as fiducial quantities, the latter defining the unit of time.

5.12 Electrons in Solids: The Band Theory

5.12.1 Origin of Energy Bands

In order to understand the principles on which the operation of semiconductor lasers is based, we must review briefly the concepts underlying the theory of electrical conduction in crystalline solids. Apart from some special cases such as the interior of a battery, electrical conduction is a manifestation of the flow of electrons. The conditions, therefore, that determine to what extent a given substance can conduct electricity have to do with the extent to which electrons are able to move freely under the action of an applied electric field.

A crystalline solid is composed of atoms (or ions) arranged in a 3-dimensional array that repeats in a regular pattern. The motion of the electrons and their quantum states are no longer determined just by the electrostatic forces within each atom individually, but rather, particularly the outer valence electrons, by the interaction with all the atoms or ions in the crystal. Instead of the atomic structure problem, where electrons are more or less attracted to a central nucleus, we now have a regular 3-dimensional array of attracting centers. To see what the quantum states of the electrons should be for such an array, let us start with just two centers initially far apart being brought together to their actual separation in the crystal. Since the two-center system is symmetric with respect to an interchange of the positions of the centers, in quantum theory it follows that the wave function representing the two-atom system must be either symmetric (unchanged) or antisymmetric (only change sign) when the electron coordinates with respect to the two centers are exchanged. Initially, when the atoms are very far apart, the energy levels computed on the basis of the two symmetries are equal, and therefore the levels are the same as in the isolated atom, except that to each energy level belong *two* possible quantum states. However, when the atoms approach each other, the energies are no longer the same for the two symmetries, and the levels are split into two close levels. If now a third atom is brought into position from a large distance, it would lead to a 3-fold exchange symmetry and a consequent splitting into three levels. By extension, if N atoms are brought into position to form a crystal, the levels are split into N levels, the widest splitting coming from nearest neighbors. Since the atomic separation determines the maximum splitting, and N for even the smallest visible piece of the crystal is extremely large, on the order of 10^{19} atoms, the result is effectively a continuous band rather than a discrete multiplet. On the basis of this band structure we can now broadly draw the essential distinctions between a conductor, an insulator, and a semiconductor.

5.12.2 Conductors and Insulators

In the lowest-energy state of the system, the electrons fill all the available states, from the lowest up to the energy band that arises from electron states in the outermost shell of the isolated atom. If the last band containing electrons is only partially filled, then there will be within that band a continuum of higher-energy states available to the electrons to go into as a result of gaining kinetic energy from an external electric field, and the crystal is a conductor. For that reason the partially filled band is called the conduction band. For example, an isolated sodium atom has one electron in its outermost 3s shell, which can accommodate, according to the Pauli principle, two electrons. The band that results from this state can therefore accommodate $2N$ electrons, whereas N sodium atoms have only half this number. Therefore, sodium is a good electrical conductor; in fact, the crystal is metallic and like all metals is a good conductor. On the other hand, a crystal is an insulator if all the bands up to a certain uppermost one, called the valence band, are completely filled in the sense of the Pauli principle, and the next higher empty band is so high in energy that no electrons can reach it by thermal agitation. In this case there are no electrons in a position to go into contiguous vacant states in response to an applied electric field, and no change in electron velocity can occur. Hence no current is produced, and the crystal is an insulator.

Finally, we have what are called semiconductors, such as pure silicon, germanium, and gallium arsenide. In these the valence band is filled like an insulator, and the band above it would be empty were it not for the circumstance that it is so close in energy to the top of the valence band that at ordinary temperatures there are appreciable numbers of electrons in it due to thermal agitation. Thus because of the thermal distribution of energy among the electrons, a semiconductor has electrons in a band that would otherwise be empty at absolute zero temperature. The vacancies left behind in the valence band by the electrons that are thermally raised to the conduction band are called *holes* and act like positive electrons. This can be made plausible by thinking of the analogy of a row of seats in a theater all occupied except one; if the person next to the vacant seat gets up and sits in it, the effect is the movement of the vacancy one seat in the opposite direction to that of the person. Clearly, the number of holes left in the valence band must equal the number of electrons in the conduction band. This number depends on the temperature according to the quantum analogue to the Maxwell–Boltzmann distribution, the *Fermi distribution*, which applies to thermal equilibrium of electrons in any system. If the probability of an electron occupying a state of energy in the interval

dE centered on the value E is defined as $F(E)dE$, then the distribution function $F(E)$ is a function of temperature of the form

$$F(E) = \frac{1}{\exp\left(\frac{E-E_F}{kT}\right)+1}, \tag{5.14}$$

where E_F is a parameter called the *Fermi energy*. Noting that at absolute zero E_F marks the energy at which $F(E)$ abruptly changes from one to zero, we see that the Fermi energy can be described as the highest level reached if all the electrons are distributed one to each of the lowest available states. From Figure 5.9 we see that in order that the number of electrons raised to the conduction band be equal to the number of holes left behind in the valence band, the Fermi energy must be assumed to be midway in the gap between the two bands. The importance of the Fermi level for us is that when a junction is formed between two types of semiconductor, the energy levels on the two sides of the junction must adjust themselves in such a way that the Fermi levels are brought into coincidence.

5.12.3 p-Type and n-Type Semiconductors

So far we have been considering ideally pure semiconductor crystals, the so-called *intrinsic* semiconductors, with impurities well below a few parts in a million. In fact, what made transistors possible and the solid-state revolution in electronics that they brought with them, are the technological advances in purifying and controlling the purity of these materials. By adding minute controlled amounts of "impurities" to the melt during the growth of the semiconductor crystals, a process called *doping*, the electrical conductivity of these semiconductors can be radically altered in

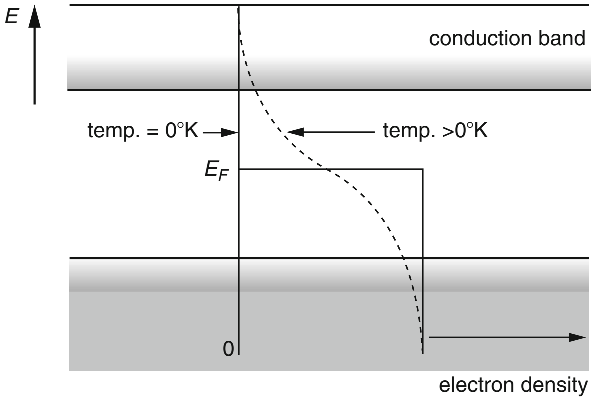

Figure 5.9 The Fermi distribution of electrons in a semiconductor

useful ways. The result of doping is what is called an *extrinsic* semiconductor, with the number of electrons exceeding that of holes (n-type), or with a preponderance of holes over electrons (p-type).

To understand better the effects of doping, we note first that elements such as silicon and germanium have a valence of four, and they crystallize in the diamond structure in which each valence electron is shared in a covalent bond with one electron from each of four nearest neighbor atoms. These covalent bonds account for all the valence electrons, and therefore at $T \rightarrow 0$ the valence band is completely filled, while the band above it, the conduction band, is empty. Suppose now that as a result of doping, some of the lattice sites in the crystal are occupied not by an atom of the host element, but by an impurity atom with a valence of five, such as arsenic. Four of these five valence electrons will be taken up in forming the four covalent bonds, leaving the fifth electron moving in the field of the remaining ion. This electron and the other such electrons belonging to impurity atoms are more weakly bound to the ions in the crystal environment than they would be in free space and therefore are in discrete states very close to the continuum of free electron states, that is, the conduction band. These discrete states are called *donor* states, because at temperatures above zero they give up electrons to the conduction band, making the crystal n-type with a high conductivity due predominantly to electron flow. The presence of the additional donor electrons puts the Fermi level closer to the conduction band.

Suppose now that the silicon or germanium crystal is doped with an impurity having a valence of three, such as aluminum or gallium. Then where an impurity atom occupies a lattice site there will be one too few electrons to satisfy the four covalent bonds. In this case, an electron from the top of the valence band supplies the missing electron to form a negative ion and leave a hole in the valence band, which, acting like a positive electron, will have weakly bound discrete states, like the mirror image of an electron in the field of a positive ion. These states will be for negative electrons slightly above the top of the valence band, and they are called *acceptor* levels, because they receive electrons from the valence band, leaving holes there to act like positive charge carriers. The resulting semiconductor is called p-type, since the predominant charge carriers responsible for conduction are positive. With fewer electrons in the valence band, the Fermi distribution must be moved lower, with E_F closer to the top of the valence band, in order again to conform with the requirement on the electron number. Figure 5.10 shows schematically the relative positions of the boundaries of the two energy bands, the impurity levels, and the Fermi levels.

5.12.4 Energy-Momentum Relationship

So far we have dealt only with the possible energy states of electrons in the crystal; but a complete dynamical description must include their momentum. This is necessary if we are to deal with the electron transitions accompanying the absorption

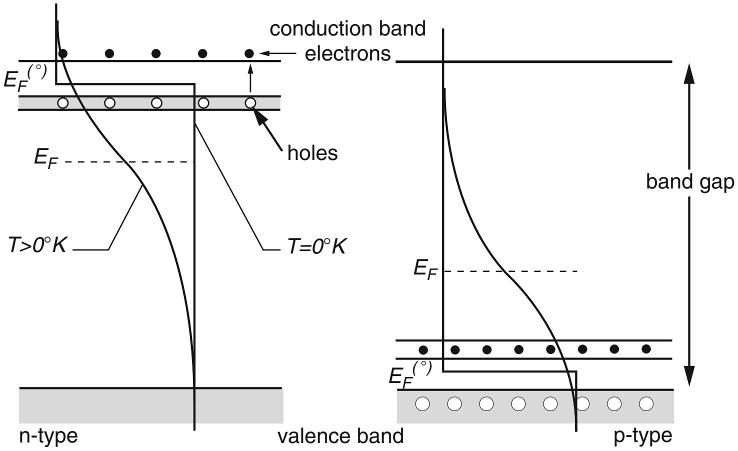

Figure 5.10 Energy bands and impurity levels in a doped semiconductor

or emission of radiation. In the case of radiative processes in atoms, conservation laws lead to certain selection rules determining which transitions are allowed and which are forbidden. Here the conservation of linear momentum between the electron making a transition and the photon absorbed or emitted will impose conditions on the crystalline properties that we must now address.

The problem of the motion of electrons acted on by a spatially periodic force such as they experience on an atomic scale from the atoms or ions in the crystal lattice is a quantum-theoretical problem. Their behavior is dominated by their wave nature, and rather than speak of the momentum of an electron, it is more useful to use the de Broglie wave vector $\mathbf{k} = m\mathbf{V}/(h/2\pi)$, whose magnitude is defined as $k = 2\pi/\lambda$. The classical (nonrelativistic) relationship between kinetic energy $E = \frac{1}{2}mV^2$ and the wave vector for a free particle is as follows:

$$E = \frac{1}{2m}\left(\frac{kh}{2\pi}\right)^2. \qquad 5.15$$

However, motion in a periodic crystalline field is totally different; in fact, even the most essential attribute of a material particle, namely its mass, is no longer a constant. The change in kinetic energy that a force imparts to an electron, that is, its "inertia," depends on its quantum state, and the concept of an "effective mass" is introduced to frame the problem where possible in Newtonian terms. The way in which the E–k relationship for a free particle is modified in an ideal crystal with a lattice spacing of a between atoms is shown schematically in Figure 5.11. We notice the band structure and the appearance of "forbidden" gaps around the points $k = n\pi/a$, where n is a whole number. These can be given an electron wave interpretation as the inability of the electron wave to propagate through the crystal with these wave numbers because of coherent reflections from the lattice sites causing destructive interference.

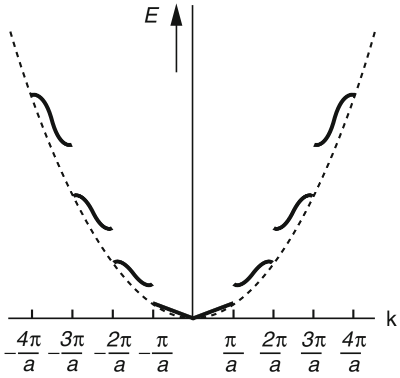

Figure 5.11 The *E–k* graph for an electron in a one-dimensional periodic field: a simple model of a crystal

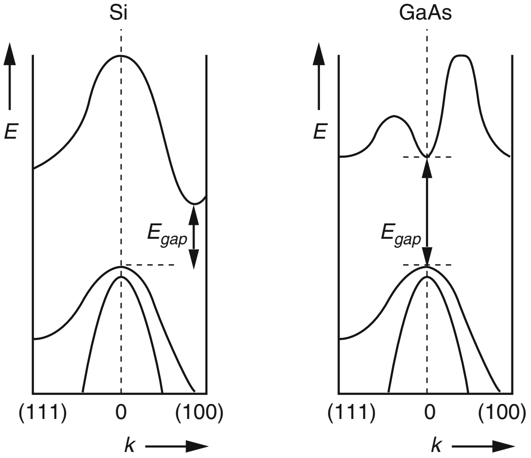

Figure 5.12 The energy–momentum graphs for silicon and gallium arsenide crystals. The indices (100) and (111) specify directions with respect to the crystal axes

In a real crystal the detailed *E–k* relationship is in general much more complicated. Figure 5.12 compares graphically the features of that relationship that are of particular relevance to us for two semiconductors: silicon and gallium arsenide. Note that the curves are for specified directions of the electron wave vector with respect to the crystal axes, since most physical properties, including electronic properties, are different in different directions in a crystal. Of particular importance is the fact that the upper boundary of the valence band for GaAs has a maximum at the same value of *k* as a minimum in the lower boundary of the conduction band, whereas for Si this is not the case. Semiconductors that are like GaAs in this respect are said to have a *direct* band-gap, while the others have an *indirect* one. We shall see later that in order that electrons may undergo radiative transitions between bands, involving the emission or absorption of photons, and do it with high probability, it is crucial that the semiconductor be a direct one.