

Free Electron Theory of Metals

Session: 2021-2022

Course code: PHY 3205

Course title: Solid State Physics I

Bachelor of Science with Honours in Physics

Jashore University of Science and Technology

Content

Band Theory: Energy Bands in Crystals; Nearly Free Electron Model and Energy Gaps; Motion of Electrons in One and Three Dimensions in a Periodic Potential; Band Theory; Bloch Theory; Kronig -Penney Model; Tight Bonding Model; LCAO Band theory.

References

Introduction to Solid State Physics – Charles Kittel

Solid State Physics – Neil W. Ashcroft, N. David Mermin

Elementary Solid State Physics: Principles and Applications – M. Ali Omar

7

Energy Bands

NEARLY FREE ELECTRON MODEL	164
Origin of the energy gap	165
Magnitude of the energy gap	167
BLOCH FUNCTIONS	167
KRONIG-PENNEY MODEL	168
WAVE EQUATION OF ELECTRON IN A PERIODIC POTENTIAL	169
Restatement of the Bloch theorem	173
Crystal momentum of an electron	173
Solution of the central equation	174
Kronig-Penney model in reciprocal space	174
Empty lattice approximation	176
Approximate solution near a zone boundary	177
NUMBER OF ORBITALS IN A BAND	180
Metals and insulators	181
SUMMARY	182
PROBLEMS	182
1. Square lattice, free electron energies	182
2. Free electron energies in reduced zone	182
3. Kronig-Penney model	182
4. Potential energy in the diamond structure	182
5. Complex wavevectors in the energy gap	183
6. Square lattice	183

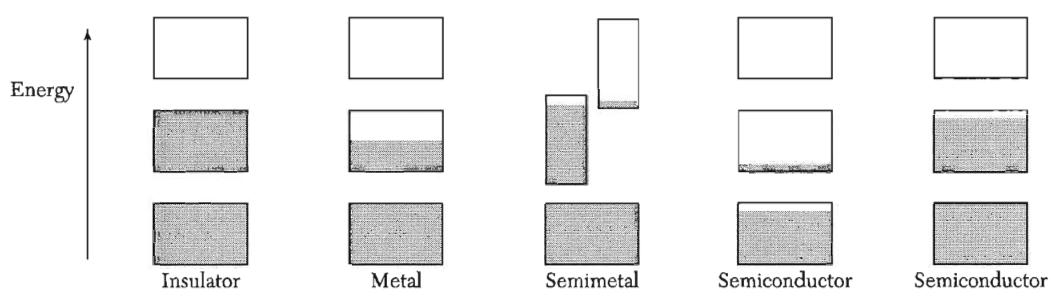


Figure 1 Schematic electron occupancy of allowed energy bands for an insulator, metal, semimetal, and semiconductor. The vertical extent of the boxes indicates the allowed energy regions; the shaded areas indicate the regions filled with electrons. In a **semimetal** (such as bismuth) one band is almost filled and another band is nearly empty at absolute zero, but a pure **semiconductor** (such as silicon) becomes an insulator at absolute zero. The left of the two semiconductors shown is at a finite temperature, with carriers excited thermally. The other semiconductor is electron-deficient because of impurities.

CHAPTER 7: ENERGY BANDS

When I started to think about it, I felt that the main problem was to explain how the electrons could sneak by all the ions in a metal. . . . By straight Fourier analysis I found to my delight that the wave differed from the plane wave of free electrons only by a periodic modulation.

F. Bloch

The free electron model of metals gives us good insight into the heat capacity, thermal conductivity, electrical conductivity, magnetic susceptibility, and electrodynamics of metals. But the model fails to help us with other large questions: the distinction between metals, semimetals, semiconductors, and insulators; the occurrence of positive values of the Hall coefficient; the relation of conduction electrons in the metal to the valence electrons of free atoms; and many transport properties, particularly magnetotransport. We need a less naïve theory, and fortunately it turns out that almost any simple attempt to improve upon the free electron model is enormously profitable.

The difference between a good conductor and a good insulator is striking. The electrical resistivity of a pure metal may be as low as 10^{-10} ohm-cm at a temperature of 1 K, apart from the possibility of superconductivity. The resistivity of a good insulator may be as high as 10^{22} ohm-cm. This range of 10^{32} may be the widest of any common physical property of solids.

Every solid contains electrons. The important question for electrical conductivity is how the electrons respond to an applied electric field. We shall see that electrons in crystals are arranged in **energy bands** (Fig. 1) separated by regions in energy for which no wavelike electron orbitals exist. Such forbidden regions are called **energy gaps** or **band gaps**, and result from the interaction of the conduction electron waves with the ion cores of the crystal.

The crystal behaves as an insulator if the allowed energy bands are either filled or empty, for then no electrons can move in an electric field. The crystal behaves as a metal if one or more bands are partly filled, say between 10 and 90 percent filled. The crystal is a **semiconductor** or a **semimetal** if one or two bands are slightly filled or slightly empty.

To understand the difference between insulators and conductors, we must extend the free electron model to take account of the periodic lattice of the solid. The possibility of a band gap is the most important new property that emerges.

We shall encounter other quite remarkable properties of electrons in crystals. For example, they respond to applied electric or magnetic fields as if the electrons were endowed with an effective mass m^* , which may be larger or smaller than the free electron mass, or may even be negative. Electrons in

crystals respond to applied fields as if endowed with negative or positive charges, $-e$ or $+e$, and herein lies the explanation of the negative and positive values of the Hall coefficient.

NEARLY FREE ELECTRON MODEL

On the free electron model the allowed energy values are distributed essentially continuously from zero to infinity. We saw in Chapter 6 that

$$\epsilon_{\mathbf{k}} = \frac{\hbar^2}{2m} (k_x^2 + k_y^2 + k_z^2) , \quad (1)$$

where, for periodic boundary conditions over a cube of side L ,

$$k_x, k_y, k_z = 0 ; \quad \pm \frac{2\pi}{L} ; \quad \pm \frac{4\pi}{L} ; \dots . \quad (2)$$

The free electron wavefunctions are of the form

$$\psi_{\mathbf{k}}(\mathbf{r}) = \exp(i\mathbf{k} \cdot \mathbf{r}) ; \quad (3)$$

they represent running waves and carry momentum $\mathbf{p} = \hbar \mathbf{k}$.

The band structure of a crystal can often be explained by the nearly free electron model for which the band electrons are treated as perturbed only weakly by the periodic potential of the ion cores. This model answers almost all the qualitative questions about the behavior of electrons in metals.

We know that Bragg reflection is a characteristic feature of wave propagation in crystals. Bragg reflection of electron waves in crystals is the cause of energy gaps. (At Bragg reflection wavelike solutions of the Schrödinger equation do not exist, as in Fig. 2.) These energy gaps are of decisive significance in determining whether a solid is an insulator or a conductor.

We explain physically the origin of energy gaps in the simple problem of a linear solid of lattice constant a . The low energy portions of the band structure

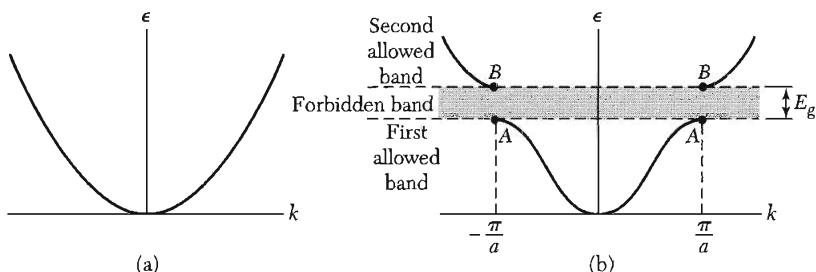


Figure 2 (a) Plot of energy ϵ versus wavevector k for a free electron. (b) Plot of energy versus wavevector for an electron in a monatomic linear lattice of lattice constant a . The energy gap E_g shown is associated with the first Bragg reflection at $k = \pm \pi/a$; other gaps are found at higher energies at $\pm n\pi/a$, for integral values of n .

are shown qualitatively in Fig. 2, in (a) for entirely free electrons and in (b) for electrons that are nearly free, but with an energy gap at $k = \pm\pi/a$. The Bragg condition $(\mathbf{k} + \mathbf{G})^2 = k^2$ for diffraction of a wave of wavevector \mathbf{k} becomes in one dimension

$$k = \pm \frac{1}{2} G = \pm n\pi/a , \quad (4)$$

where $G = 2\pi n/a$ is a reciprocal lattice vector and n is an integer. The first reflections and the first energy gap occur at $k = \pm\pi/a$. The region in \mathbf{k} space between $-\pi/a$ and π/a is the **first Brillouin zone** of this lattice. Other energy gaps occur for other values of the integer n .

The wavefunctions at $k = \pm\pi/a$ are not the traveling waves $\exp(i\pi x/a)$ or $\exp(-i\pi x/a)$ of free electrons. At these special values of k the wavefunctions are made up of equal parts of waves traveling to the right and to the left. When the Bragg reflection condition $k = \pm\pi/a$ is satisfied by the wavevector, a wave traveling to the right is Bragg-reflected to travel to the left, and vice versa. Each subsequent Bragg reflection will reverse the direction of travel of the wave. A wave that travels neither to the right nor to the left is a standing wave: it doesn't go anywhere.

The time-independent state is represented by standing waves. We can form two different standing waves from the two traveling waves

$$\exp(\pm i\pi x/a) = \cos(\pi x/a) \pm i \sin(\pi x/a),$$

so that the standing waves are

$$\begin{aligned} \psi(+) &= \exp(i\pi x/a) + \exp(-i\pi x/a) = 2 \cos(\pi x/a) ; \\ \psi(-) &= \exp(i\pi x/a) - \exp(-i\pi x/a) = 2i \sin(\pi x/a) . \end{aligned} \quad (5)$$

The standing waves are labeled (+) or (-) according to whether or not they change sign when $-x$ is substituted for x . Both standing waves are composed of equal parts of right- and left-directed traveling waves.

Origin of the Energy Gap

The two standing waves $\psi(+)$ and $\psi(-)$ pile up electrons at different regions, and therefore the two waves have different values of the potential energy in the field of the ions of the lattice. This is the origin of the energy gap. The probability density ρ of a particle is $\psi^* \psi = |\psi|^2$. For a pure traveling wave $\exp(ikx)$, we have $\rho = \exp(-ikx) \exp(ikx) = 1$, so that the charge density is constant. The charge density is not constant for linear combinations of plane waves. Consider the standing wave $\psi(+)$ in (5); for this we have

$$\rho(+) = |\psi(+)|^2 \propto \cos^2 \pi x/a .$$

This function piles up electrons (negative charge) on the positive ions centered at $x = 0, a, 2a, \dots$ in Fig. 3, where the potential energy is lowest.

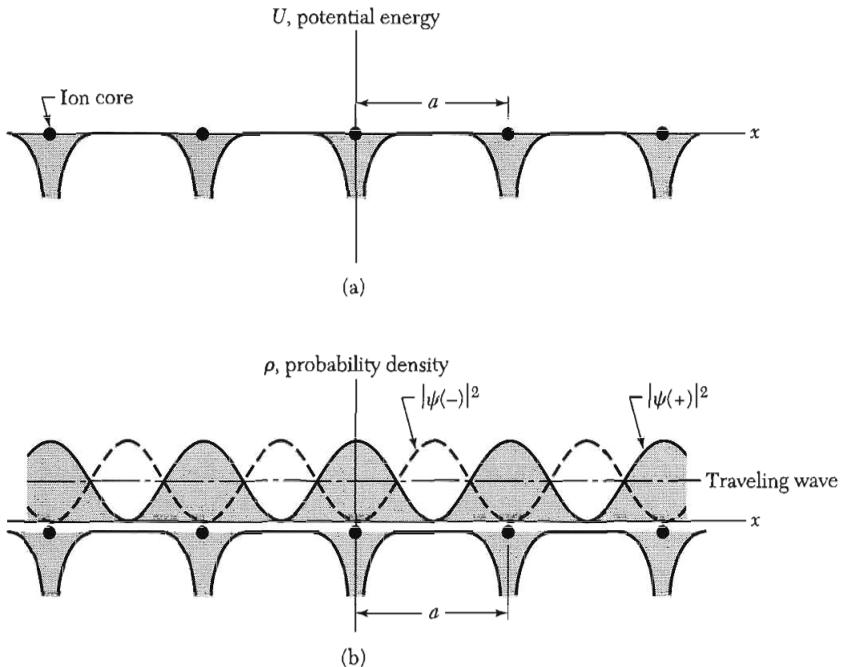


Figure 3 (a) Variation of potential energy of a conduction electron in the field of the ion cores of a linear lattice. (b) Distribution of probability density ρ in the lattice for $|\psi(-)|^2 \propto \sin^2 \pi x/a$; $|\psi(+)|^2 \propto \cos^2 \pi x/a$; and for a traveling wave. The wavefunction $\psi(+)$ piles up electronic charge on the cores of the positive ions, thereby lowering the potential energy in comparison with the average potential energy seen by a traveling wave. The wavefunction $\psi(-)$ piles up charge in the region between the ions, thereby raising the potential energy in comparison with that seen by a traveling wave. This figure is the key to understanding the origin of the energy gap.

Figure 3a pictures the variation of the electrostatic potential energy of a conduction electron in the field of the positive ion cores. The ion cores bear a net positive charge because the atoms are ionized in the metal, with the valence electrons taken off to form the conduction band. The potential energy of an electron in the field of a positive ion is negative, so that the force between them is attractive.

For the other standing wave $\psi(-)$ the probability density is

$$\rho(-) = |\psi(-)|^2 \propto \sin^2 \pi x/a ,$$

which concentrates electrons away from the ion cores. In Fig. 3b we show the electron concentration for the standing waves $\psi(+)$, $\psi(-)$, and for a traveling wave.

When we calculate the average or expectation values of the potential energy over these three charge distributions, we find that the potential energy of $\rho(+)$ is lower than that of the traveling wave, whereas the potential energy of $\rho(-)$ is higher than the traveling wave. We have an energy gap of width E_g if

the energies of $\rho(-)$ and $\rho(+)$ differ by E_g . Just below the energy gap at points A in Fig. 2 the wavefunction is $\psi(+)$, and just above the gap at points B the wavefunction is $\psi(-)$.

Magnitude of the Energy Gap

The wavefunctions at the Brillouin zone boundary $k = \pi/a$ are $\sqrt{2} \cos \pi x/a$ and $\sqrt{2} \sin \pi x/a$, normalized over unit length of line. Let us suppose that the potential energy of an electron in the crystal at point x is

$$U(x) = U \cos 2\pi x/a .$$

The first-order energy difference between the two standing wave states is

$$\begin{aligned} E_g &= \int_0^1 dx U(x) [|\psi(+)|^2 - |\psi(-)|^2] \\ &= 2 \int dx U \cos(2\pi x/a) (\cos^2 \pi x/a - \sin^2 \pi x/a) = U . \end{aligned} \quad (6)$$

We see that the gap is equal to the Fourier component of the crystal potential.

BLOCH FUNCTIONS

F. Bloch proved the important theorem that the solutions of the Schrödinger equation for a periodic potential must be of a special form:

$$\boxed{\psi_{\mathbf{k}}(\mathbf{r}) = u_{\mathbf{k}}(\mathbf{r}) \exp(i\mathbf{k} \cdot \mathbf{r}) ,} \quad (7)$$

where $u_{\mathbf{k}}(\mathbf{r})$ has the period of the crystal lattice with $u_{\mathbf{k}}(\mathbf{r}) = u_{\mathbf{k}}(\mathbf{r} + \mathbf{T})$. Here \mathbf{T} is a translation vector of the lattice. The result (7) expresses the Bloch theorem:

The eigenfunctions of the wave equation for a periodic potential are the product of a plane wave $\exp(i\mathbf{k} \cdot \mathbf{r})$ times a function $u_{\mathbf{k}}(\mathbf{r})$ with the periodicity of the crystal lattice.

A one-electron wavefunction of the form (7) is called a Bloch function and can be decomposed into a sum of traveling waves, as we see later. Bloch functions can be assembled into wave packets to represent electrons that propagate freely through the potential field of the ion cores.

We give now a restricted proof of the Bloch theorem, valid when ψ_k is nondegenerate; that is, when there is no other wavefunction with the same energy and wavevector as ψ_k . The general case will be treated later. We consider N identical lattice points on a ring of length Na . The potential energy is periodic in a , with $U(x) = U(x + sa)$, where s is an integer.

Let us be guided by the symmetry of the ring to look for solutions of the wave equation such that

$$\psi(x + a) = C\psi(x) , \quad (8)$$

where C is a constant. Then, on going once around the ring,

$$\psi(x + Na) = \psi(x) = C^N \psi(x) ,$$

because $\psi(x)$ must be single-valued. It follows that C is one of the N roots of unity, or

$$C = \exp(i2\pi s/N) ; \quad s = 0, 1, 2, \dots, N - 1 . \quad (9)$$

We use (9) to see that

$$\psi(x) = u_k(x) \exp(i2\pi s x / Na) \quad (10)$$

satisfies (8), provided that $u_k(x)$ has the periodicity a , so that $u_k(x) = u_k(x + a)$. This is the Bloch result (7).

KRONIG-PENNEY MODEL

A periodic potential for which the wave equation can be solved in terms of elementary functions is the square-well array of Fig. 4. The wave equation is

$$-\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} + U(x)\psi = \epsilon\psi , \quad (11)$$

where $U(x)$ is the potential energy and ϵ is the energy eigenvalue.

In the region $0 < x < a$ in which $U = 0$, the eigenfunction is a linear combination,

$$\psi = Ae^{ikx} + Be^{-ikx} , \quad (12)$$

of plane waves traveling to the right and to the left, with energy

$$\epsilon = \hbar^2 K^2 / 2m . \quad (13)$$

In the region $-b < x < 0$ within the barrier the solution is of the form

$$\psi = Ce^{Qx} + De^{-Qx} , \quad (14)$$

with

$$U_0 - \epsilon = \hbar^2 Q^2 / 2m . \quad (15)$$

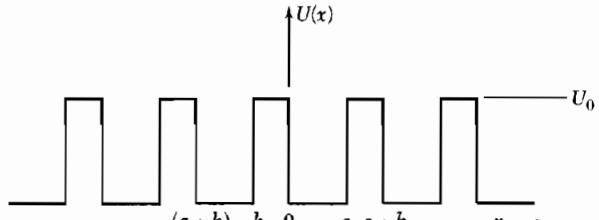


Figure 4 Square-well periodic potential as introduced by Kronig and Penney.

We want the complete solution to have the Bloch form (7). Thus the solution in the region $a < x < a + b$ must be related to the solution (14) in the region $-b < x < 0$ by the Bloch theorem:

$$\psi(a < x < a + b) = \psi(-b < x < 0) e^{ik(a+b)} , \quad (16)$$

which serves to define the wavevector k used as an index to label the solution.

The constants A, B, C, D are chosen so that ψ and $d\psi/dx$ are continuous at $x = 0$ and $x = a$. These are the usual quantum mechanical boundary conditions in problems that involve square potential wells. At $x = 0$,

$$A + B = C + D ; \quad (17)$$

$$iK(A - B) = Q(C - D) , \quad (18)$$

with Q from (14). At $x = a$, with the use of (16) for $\psi(a)$ under the barrier in terms of $\psi(-b)$,

$$Ae^{iKa} + Be^{-iKa} = (Ce^{-Qb} + De^{Qb}) e^{ik(a+b)} ; \quad (19)$$

$$iK(Ae^{iKa} - Be^{-iKa}) = Q(Ce^{-Qb} - De^{Qb}) e^{ik(a+b)} . \quad (20)$$

The four equations (17) to (20) have a solution only if the determinant of the coefficients of A, B, C, D vanishes, yielding

$$[(Q^2 - K^2)/2QK] \sinh Qb \sin Ka + \cosh Qb \cos Ka = \cos k(a + b) . \quad (21a)$$

It is rather tedious to obtain this equation.

The result is simplified if we represent the potential by the periodic delta function obtained when we pass to the limit $b = 0$ and $U_0 = \infty$ in such a way that $Q^2ba/2 = P$, a finite quantity. In this limit $Q \gg K$ and $Qb \ll 1$. Then (21a) reduces to

$$(P/Ka) \sin Ka + \cos Ka = \cos ka . \quad (21b)$$

The ranges of K for which this equation has solutions are plotted in Fig. 5, for the case $P = 3\pi/2$. The corresponding values of the energy are plotted in Fig. 6. Note the energy gaps at the zone boundaries. The wavevector k of the Bloch function is the important index, not the K in (12), which is related to the energy by (13). A treatment of this problem in wavevector space is given later in this chapter.

WAVE EQUATION OF ELECTRON IN A PERIODIC POTENTIAL

We considered in Fig. 3 the approximate form we expect for the solution of the Schrödinger equation if the wavevector is at a zone boundary, as at $k = \pm\pi/a$. We treat in detail the wave equation for a general potential, at general

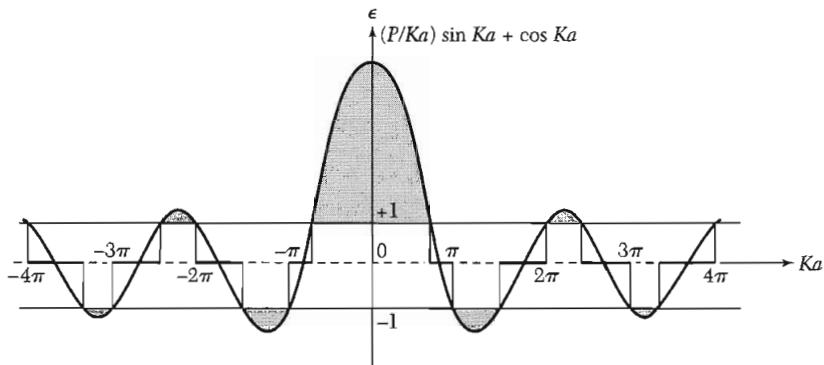


Figure 5 Plot of the function $(P/Ka) \sin Ka + \cos Ka$, for $P = 3\pi/2$. The allowed values of the energy e are given by those ranges of $Ka = (2me/\hbar^2)a$ for which the function lies between ± 1 . For other values of the energy there are no traveling wave or Bloch-like solutions to the wave equation, so that forbidden gaps in the energy spectrum are formed.

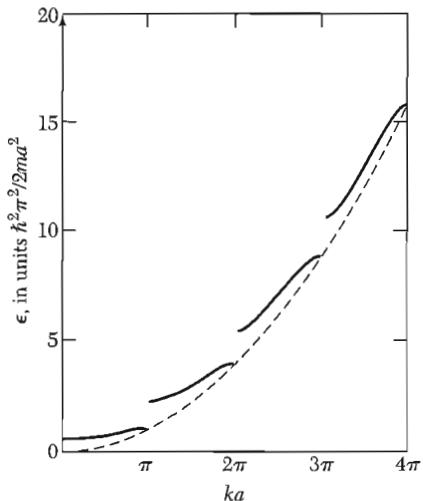


Figure 6 Plot of energy vs. wavenumber for the Kronig-Penney potential, with $P = 3\pi/2$. Notice the energy gaps at $ka = \pi, 2\pi, 3\pi \dots$

values of k . Let $U(x)$ denote the potential energy of an electron in a linear lattice of lattice constant a . We know that the potential energy is invariant under a crystal lattice translation: $U(x) = U(x + a)$. A function invariant under a crystal lattice translation may be expanded as a Fourier series in the reciprocal lattice vectors G . We write the Fourier series for the potential energy as

$$U(x) = \sum_G U_G e^{iGx}. \quad (22)$$

The values of the coefficients U_G for actual crystal potentials tend to decrease rapidly with increasing magnitude of G . For a bare coulomb potential U_G decreases as $1/G^2$.

We want the potential energy $U(x)$ to be a real function:

$$U(x) = \sum_{G>0} U_G (e^{iGx} + e^{-iGx}) = 2 \sum_{G>0} U_G \cos Gx . \quad (23)$$

For convenience we have assumed that the crystal is symmetric about $x = 0$ and that $U_0 = 0$.

The wave equation of an electron in the crystal is $\mathcal{H}\psi = \epsilon\psi$, where \mathcal{H} is the hamiltonian and ϵ is the energy eigenvalue. The solutions ψ are called eigenfunctions or orbitals or Bloch functions. Explicitly, the wave equation is

$$\left(\frac{1}{2m} p^2 + U(x) \right) \psi(x) = \left(\frac{1}{2m} p^2 + \sum_G U_G e^{iGx} \right) \psi(x) = \epsilon \psi(x) . \quad (24)$$

Equation (24) is written in the one-electron approximation in which the orbital $\psi(x)$ describes the motion of one electron in the potential of the ion cores and in the average potential of the other conduction electrons.

The wavefunction $\psi(x)$ may be expressed as a Fourier series summed over all values of the wavevector permitted by the boundary conditions, so that

$$\psi = \sum_k C(k) e^{ikx} , \quad (25)$$

where k is real. (We could equally well write the index k as a subscript on C , as in C_k)

The set of values of k has the form $2\pi n/L$, because these values satisfy periodic boundary conditions over length L . Here n is any integer, positive or negative. We do not assume, nor is it generally true, that $\psi(x)$ itself is periodic in the fundamental lattice translation a . The translational properties of $\psi(x)$ are determined by the Bloch theorem (7).

Not all wavevectors of the set $2\pi n/L$ enter the Fourier expansion of any one Bloch function. If one particular wavevector k is contained in a ψ , then all other wavevectors in the Fourier expansion of this ψ will have the form $k + G$, where G is any reciprocal lattice vector. We prove this result in (29) below.

We can label a wavefunction ψ that contains a component k as ψ_k or, equally well, as ψ_{k+G} , because if k enters the Fourier expansion then $k + G$ may enter. The wavevectors $k + G$ running over G are a restricted subset of the set $2\pi n/L$, as shown in Fig. 7.

We shall usually choose as a label for the Bloch function that k which lies within the first Brillouin zone. When other conventions are used, we shall say so. This situation differs from the phonon problem for a monatomic lattice where there are no components of the ion motion outside the first zone. The electron problem is like the x-ray diffraction problem because like the electron wavefunction the electromagnetic field exists everywhere within the crystal and not only at the ions.

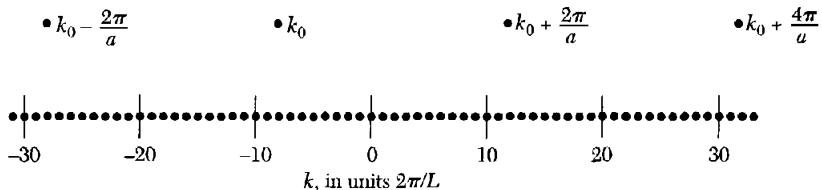


Figure 7 The lower points represent values of the wavevector $k = 2\pi n/L$ allowed by the periodic boundary condition on the wavefunction over a ring of circumference L composed of 20 primitive cells. The allowed values continue to $\pm\infty$. The upper points represent the first few wavevectors which may enter into the Fourier expansion of a wavefunction $\psi(x)$, starting from a particular wavevector $k = k_0 = -8(2\pi/L)$. The shortest reciprocal lattice vector is $2\pi/a = 20(2\pi/L)$.

To solve the wave equation, substitute (25) in (24) to obtain a set of linear algebraic equations for the Fourier coefficients. The kinetic energy term is

$$\frac{1}{2m} p^2 \psi(x) = \frac{1}{2m} \left(-i\hbar \frac{d}{dx} \right)^2 \psi(x) = -\frac{\hbar^2}{2m} \frac{d^2 \psi}{dx^2} = \frac{\hbar^2}{2m} \sum_k k^2 C(k) e^{ikx} ;$$

and the potential energy term is

$$\left(\sum_G U_G e^{iGx} \right) \psi(x) = \sum_G \sum_k U_G e^{iGx} C(k) e^{ikx} .$$

The wave equation is obtained as the sum:

$$\sum_k \frac{\hbar^2}{2m} k^2 C(k) e^{ikx} + \sum_G \sum_k U_G C(k) e^{i(k+G)x} = \epsilon \sum_k C(k) e^{ikx} . \quad (26)$$

Each Fourier component must have the same coefficient on both sides of the equation. Thus we have the **central equation**

$$(\lambda_k - \epsilon) C(k) + \sum_G U_G C(k - G) = 0 . \quad (27)$$

with the notation

$$\lambda_k = \hbar^2 k^2 / 2m . \quad (28)$$

Equation (27) is a useful form of the wave equation in a periodic lattice, although unfamiliar because a set of algebraic equations has taken the place of the usual differential equation (24). The set appears unpleasant and formidable because there are, in principle, an infinite number of $C(k - G)$ to be determined. In practice a small number will often suffice, perhaps two or four. It takes some experience to appreciate the practical advantages of the algebraic approach.

Restatement of the Bloch Theorem

Once we determine the C 's from (27), the wavefunction (25) is given as

$$\psi_k(x) = \sum_G C(k - G) e^{i(k-G)x} , \quad (29)$$

which may be rearranged as

$$\psi_k(x) = \left(\sum_G C(k - G) e^{-iGx} \right) e^{ikx} = e^{ikx} u_k(x) ,$$

with the definition

$$u_k(x) \equiv \sum_G C(k - G) e^{-iGx} .$$

Because $u_k(x)$ is a Fourier series over the reciprocal lattice vectors, it is invariant under a crystal lattice translation T , so that $u_k(x) = u_k(x + T)$. We verify this directly by evaluating $u_k(x + T)$:

$$u_k(x + T) = \sum_G C(k - G) e^{-iG(x+T)} = e^{-iGT} \left[\sum_G C(k - G) e^{-iGx} \right] = e^{-iGT} u_k(x) .$$

Because $\exp(-iGT) = 1$ by (2.17), it follows that $u_k(x + T) = u_k(x)$, thereby establishing the periodicity of u_k . This is an alternate and exact proof of the Bloch theorem and is valid even when the ψ_k are degenerate.

Crystal Momentum of an Electron

What is the significance of the wavevector \mathbf{k} used to label the Bloch function? It has several properties:

- Under a crystal lattice translation which carries \mathbf{r} to $\mathbf{r} + \mathbf{T}$ we have

$$\psi_{\mathbf{k}}(\mathbf{r} + \mathbf{T}) = e^{i\mathbf{k} \cdot \mathbf{T}} e^{i\mathbf{k} \cdot \mathbf{r}} u_{\mathbf{k}}(\mathbf{r} + \mathbf{T}) = e^{i\mathbf{k} \cdot \mathbf{T}} \psi_{\mathbf{k}}(\mathbf{r}) , \quad (30)$$

because $u_{\mathbf{k}}(\mathbf{r} + \mathbf{T}) = u_{\mathbf{k}}(\mathbf{r})$. Thus $\exp(i\mathbf{k} \cdot \mathbf{T})$ is the phase factor by which a Bloch function is multiplied when we make a crystal lattice translation \mathbf{T} .

- If the lattice potential vanishes, the central equation (27) reduces to $(\lambda_{\mathbf{k}} - \epsilon)C(\mathbf{k}) = 0$, so that all $C(\mathbf{k} - \mathbf{G})$ are zero except $C(\mathbf{k})$, and thus $u_{\mathbf{k}}(\mathbf{r})$ is constant. We have $\psi_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k} \cdot \mathbf{r}}$, just as for a free electron. (This assumes we have had the foresight to pick the “right” \mathbf{k} as the label. For many purposes other choices of \mathbf{k} , differing by a reciprocal lattice vector, will be more convenient.)
- The quantity \mathbf{k} enters in the conservation laws that govern collision processes in crystals. (The conservation laws are really selection rules for transitions.) Thus $i\mathbf{k}$ is called the **crystal momentum** of an electron. If an electron \mathbf{k} absorbs in a collision a phonon of wavevector \mathbf{q} , the selection rule is $\mathbf{k} + \mathbf{q} = \mathbf{k}' + \mathbf{G}$. In this process the electron is scattered from a state \mathbf{k} to a state \mathbf{k}' , with \mathbf{G} a reciprocal lattice vector. Any arbitrariness in labeling the Bloch functions can be absorbed in the \mathbf{G} without changing the physics of the process.

Solution of the Central Equation

The central equation (27),

$$(\lambda_k - \epsilon)C(k) + \sum_G U_G C(k - G) = 0 , \quad (31)$$

represents a set of simultaneous linear equations that connect the coefficients $C(k - G)$ for all reciprocal lattice vectors G . It is a set because there are as many equations as there are coefficients C . These equations are consistent if the determinant of the coefficients vanishes.

Let us write out the equations for an explicit problem. We let g denote the shortest G . We suppose that the potential energy $U(x)$ contains only a single Fourier component $U_g = U_{-g}$, denoted by U . Then a block of the determinant of the coefficients is given by:

$$\begin{array}{ccccc} \lambda_{k-2g} - \epsilon & U & 0 & 0 & 0 \\ U & \lambda_{k-g} - \epsilon & U & 0 & 0 \\ 0 & U & \lambda_k - \epsilon & U & 0 \\ 0 & 0 & U & \lambda_{k+g} - \epsilon & U \\ 0 & 0 & 0 & U & \lambda_{k+2g} - \epsilon \end{array} . \quad (32)$$

To see this, write out five successive equations of the set (31). The determinant in principle is infinite in extent, but it will often be sufficient to set equal to zero the portion we have shown.

At a given k , each root ϵ or ϵ_k lies on a different energy band, except in case of coincidence. The solution of the determinant (32) gives a set of energy eigenvalues ϵ_{nk} , where n is an index for ordering the energies and k is the wavevector that labels C_k .

Most often k will be taken in the first zone, to reduce possible confusion in the labeling. If we chose a k different from the original by some reciprocal lattice vector, we would have obtained the same set of equations in a different order—but having the same energy spectrum.

Kronig-Penney Model in Reciprocal Space

As an example of the use of the central equation (31) for a problem that is exactly solvable, we use the Kronig-Penney model of a periodic delta-function potential:

$$U(x) = 2 \sum_{G>0} U_G \cos Gx = Aa \sum_s \delta(x - sa) , \quad (33)$$

where A is a constant and a the lattice spacing. The sum is over all integers s between 0 and $1/a$. The boundary conditions are periodic over a ring of unit

length, which means over $1/a$ atoms. Thus the Fourier coefficients of the potential are

$$\begin{aligned} U_G &= \int_0^1 dx U(x) \cos Gx = Aa \sum_s \int_0^1 dx \delta(x - sa) \cos Gx \\ &= Aa \sum_s \cos Gsa = A . \end{aligned} \quad (34)$$

All U_G are equal for the delta-function potential.

We write the central equation with k as the Bloch index. Thus (31) becomes

$$(\lambda_k - \epsilon)C(k) + A \sum_n C(k - 2\pi n/a) = 0 , \quad (35)$$

where $\lambda_k \equiv \hbar^2 k^2 / 2m$ and the sum is over all integers n . We want to solve (35) for $\epsilon(k)$.

We define

$$f(k) = \sum_n C(k - 2\pi n/a) , \quad (36)$$

so that (35) becomes

$$C(k) = - \frac{(2mA/\hbar^2)f(k)}{k^2 - (2m\epsilon/\hbar^2)} . \quad (37)$$

Because the sum (36) is over all coefficients C , we have, for any n ,

$$f(k) = f(k - 2\pi n/a) . \quad (38)$$

This relation lets us write

$$C(k - 2\pi n/a) = - (2mA/\hbar^2)f(k)[(k - 2\pi n/a)^2 - 2m\epsilon/\hbar^2]^{-1} . \quad (39)$$

We sum both sides over all n to obtain, using (36) and cancelling $f(k)$ from both sides,

$$(\hbar^2/2mA) = - \sum_n [(k - 2\pi n/a)^2 - (2m\epsilon/\hbar^2)]^{-1} . \quad (40)$$

The sum can be carried out with the help of the standard relation

$$\operatorname{ctn} x = \sum_n \frac{1}{n\pi + x} . \quad (41)$$

After trigonometric manipulations in which we use relations for the difference of two cotangents and the product of two sines, the sum in (40) becomes

$$\frac{a^2 \sin Ka}{4Ka(\cos ka - \cos Ka)} , \quad (42)$$

where we write $K^2 = 2m\epsilon/\hbar^2$ as in (13).

The final result for (40) is

$$(mAa^2/2\hbar^2)(Ka)^{-1} \sin Ka + \cos Ka = \cos ka , \quad (43)$$

which agrees with the Kronig-Penney result (21b) with P written for $mAa^2/2\hbar^2$.

Empty Lattice Approximation

Actual band structures are usually exhibited as plots of energy versus wavevector in the first Brillouin zone. When wavevectors happen to be given outside the first zone, they are carried back into the first zone by subtracting a suitable reciprocal lattice vector. Such a translation can always be found. The operation is helpful in visualization.

When band energies are approximated fairly well by free electron energies $\epsilon_{\mathbf{k}} = \hbar^2 \mathbf{k}^2 / 2m$, it is advisable to start a calculation by carrying the free electron energies back into the first zone. The procedure is simple enough once one gets the hang of it. We look for a \mathbf{G} such that a \mathbf{k}' in the first zone satisfies

$$\mathbf{k}' + \mathbf{G} = \mathbf{k} ,$$

where \mathbf{k} is unrestricted and is the true free electron wavevector in the empty lattice. (Once the plane wave is modulated by the lattice, there is no single "true" wavevector for the state ψ .)

If we drop the prime on \mathbf{k}' as unnecessary baggage, the free electron energy can always be written as

$$\begin{aligned} \epsilon(k_x, k_y, k_z) &= (\hbar^2/2m)(\mathbf{k} + \mathbf{G})^2 \\ &= (\hbar^2/2m)[(k_x + G_x)^2 + (k_y + G_y)^2 + (k_z + G_z)^2] , \end{aligned}$$

with \mathbf{k} in the first zone and \mathbf{G} allowed to run over the appropriate reciprocal lattice points.

We consider as an example the low-lying free electron bands of a simple cubic lattice. Suppose we want to exhibit the energy as a function of \mathbf{k} in the [100] direction. For convenience, choose units such that $\hbar^2/2m = 1$. We show several low-lying bands in this empty lattice approximation with their energies $\epsilon(000)$ at $\mathbf{k} = 0$ and $\epsilon(k_x 00)$ along the k_x axis in the first zone:

Band	$Ga/2\pi$	$\epsilon(000)$	$\epsilon(k_x 00)$
1	000	0	k_x^2
2,3	100,100	$(2\pi/a)^2$	$(k_x \pm 2\pi/a)^2$
4,5,6,7	010,0\bar{1}0,001,00\bar{1}	$(2\pi/a)^2$	$k_x^2 + (2\pi/a)^2$
8,9,10,11	110,101,1\bar{1}0,10\bar{1}	$2(2\pi/a)^2$	$(k_x + 2\pi/a)^2 + (2\pi/a)^2$
12,13,14,15	\bar{1}10,\bar{1}01,1\bar{1}0,\bar{1}0\bar{1}	$2(2\pi/a)^2$	$(k_x - 2\pi/a)^2 + (2\pi/a)^2$
16,17,18,19	011,0\bar{1}\bar{1},01\bar{1},0\bar{1}\bar{1}	$2(2\pi/a)^2$	$k_x^2 + 2(2\pi/a)^2$

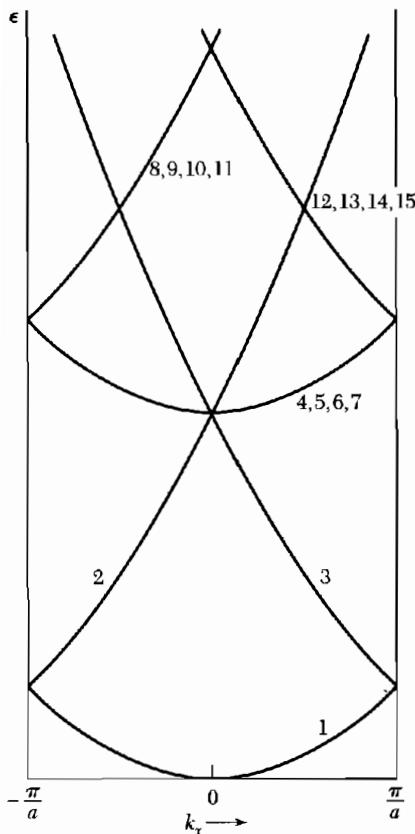


Figure 8 Low-lying free electron energy bands of the empty *sc* lattice, as transformed to the first Brillouin zone and plotted vs. $(k_x, 0)$. The free electron energy is $\hbar^2(\mathbf{k} + \mathbf{G})^2/2m$, where the \mathbf{G} 's are given in the second column of the table. The bold curves are in the first Brillouin zone, with $-\pi/a \leq k_x \leq \pi/a$. Energy bands drawn in this way are said to be in the reduced zone scheme.

These free electron bands are plotted in Fig. 8. It is a good exercise to plot the same bands for \mathbf{k} parallel to the [111] direction of wavevector space.

Approximate Solution Near a Zone Boundary

We suppose that the Fourier components U_G of the potential energy are small in comparison with the kinetic energy of a free electron at the zone boundary. We first consider a wavevector exactly at the zone boundary at $\frac{1}{2}\mathbf{G}$, that is, at π/a . Here

$$k^2 = (\frac{1}{2}\mathbf{G})^2 ; \quad (k - \mathbf{G})^2 = (\frac{1}{2}\mathbf{G} - \mathbf{G})^2 = (\frac{1}{2}\mathbf{G})^2 ,$$

so that at the zone boundary the kinetic energy of the two component waves $k = \pm\frac{1}{2}\mathbf{G}$ are equal.

If $C(\frac{1}{2}\mathbf{G})$ is an important coefficient in the orbital (29) at the zone boundary, then $C(-\frac{1}{2}\mathbf{G})$ is also an important coefficient. This result also follows from the discussion of (5). We retain only those equations in the central equation that contain both coefficients $C(\frac{1}{2}\mathbf{G})$ and $C(-\frac{1}{2}\mathbf{G})$, and neglect all other coefficients.

One equation of (31) becomes, with $k = \frac{1}{2}G$ and $\lambda = \hbar^2(\frac{1}{2}G)^2/2m$,

$$(\lambda - \epsilon)C(\frac{1}{2}G) + UC(-\frac{1}{2}G) = 0 . \quad (44)$$

Another equation of (31) becomes, with $k = \frac{1}{2}G$,

$$(\lambda - \epsilon)C(-\frac{1}{2}G) + UC(\frac{1}{2}G) = 0 . \quad (45)$$

These two equations have nontrivial solutions for the two coefficients if the energy ϵ satisfies

$$\begin{vmatrix} \lambda - \epsilon & U \\ U & \lambda - \epsilon \end{vmatrix} = 0 , \quad (46)$$

whence

$$(\lambda - \epsilon)^2 = U^2 ; \quad \epsilon = \lambda \pm U = \frac{\hbar^2}{2m}(\frac{1}{2}G)^2 \pm U . \quad (47)$$

The energy has two roots, one lower than the free electron kinetic energy by U , and one higher by U . Thus the potential energy $2U \cos Gx$ has created an energy gap $2U$ at the zone boundary.

The ratio of the C 's may be found from either (44) or (45):

$$\frac{C(-\frac{1}{2}G)}{C(\frac{1}{2}G)} = \frac{\epsilon - \lambda}{U} = \pm 1 , \quad (48)$$

where the last step uses (47). Thus the Fourier expansion of $\psi(x)$ at the zone boundary has the two solutions

$$\psi(x) = \exp(iGx/2) \pm \exp(-iGx/2) .$$

These orbitals are identical to (5).

One solution gives the wavefunction at the bottom of the energy gap; the other gives the wavefunction at the top of the gap. Which solution has the lower energy depends on the sign of U .

We now solve for orbitals with wavevector k near the zone boundary $\frac{1}{2}G$. We use the same two-component approximation, now with a wavefunction of the form

$$\psi(x) = C(k) e^{ikx} + C(k - G) e^{i(k-G)x} . \quad (49)$$

As directed by the central equation (31), we solve the pair of equations

$$(\lambda_k - \epsilon)C(k) + UC(k - G) = 0 ;$$

$$(\lambda_{k-G} - \epsilon)C(k - G) + UC(k) = 0 ,$$

with λ_k defined as $\hbar^2 k^2 / 2m$. These equations have a solution if the energy ϵ satisfies

$$\begin{vmatrix} \lambda_k - \epsilon & U \\ U & \lambda_{k-G} - \epsilon \end{vmatrix} = 0 ,$$

whence $\epsilon^2 - \epsilon(\lambda_{k-G} + \lambda_k) + \lambda_{k-G}\lambda_k - U^2 = 0$.

The energy has two roots:

$$\epsilon = \frac{1}{2}(\lambda_{k-G} + \lambda_k) \pm [\frac{1}{4}(\lambda_{k-G} - \lambda_k)^2 + U^2]^{1/2}, \quad (50)$$

and each root describes an energy band, plotted in Fig. 9. It is convenient to expand the energy in terms of a quantity \tilde{K} (the mark over the K is called a tilde), which measures the difference $\tilde{K} \equiv k - \frac{1}{2}G$ in wavevector between k and the zone boundary:

$$\begin{aligned} \epsilon_{\tilde{K}} &= (\hbar^2/2m)(\frac{1}{4}G^2 + \tilde{K}^2) \pm [4\lambda(\hbar^2\tilde{K}^2/2m) + U^2]^{1/2} \\ &\approx (\hbar^2/2m)\frac{1}{4}G^2 + \tilde{K}^2 \pm U[1 + 2(\lambda/U^2)(\hbar^2\tilde{K}^2/2m)], \end{aligned} \quad (51)$$

in the region $\hbar^2 G \tilde{K} / 2m \ll |U|$. Here $\lambda = (\hbar^2/2m)(\frac{1}{2}G)^2$ as before.

Writing the two zone boundary roots of (47) as $\epsilon(\pm)$, we may write (51) as

$$\epsilon_{\tilde{K}}(\pm) = \epsilon(\pm) + \frac{\hbar^2 \tilde{K}^2}{2m} \left(1 \pm \frac{2\lambda}{U}\right). \quad (52)$$

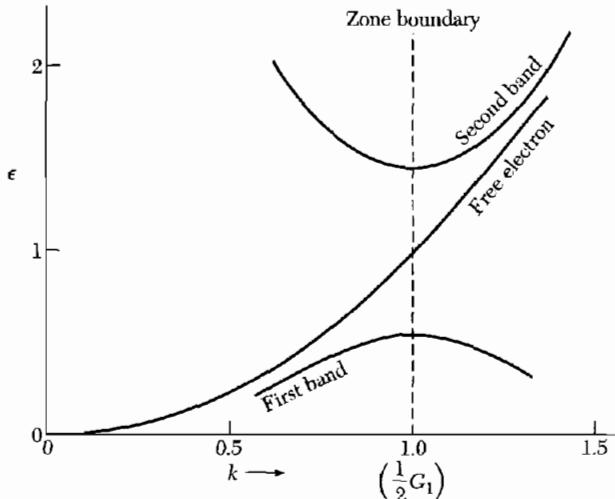


Figure 9 Solutions of (50) in the periodic zone scheme, in the region near a boundary of the first Brillouin zone. The units are such that $U = -0.45$, $G = 2$, and $\hbar^2/m = 1$. The free electron curve is drawn for comparison. The energy gap at the zone boundary is 0.90. The value of U has deliberately been chosen large for this illustration, too large for the two-term approximation to be accurate.

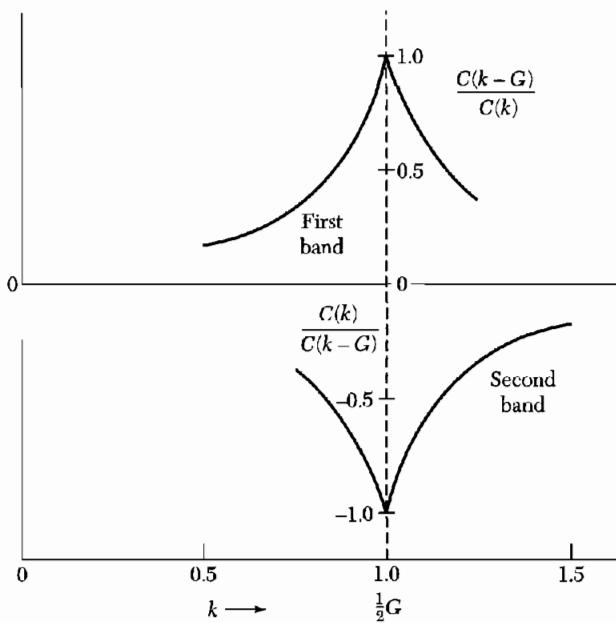


Figure 10 Ratio of the coefficients in $\psi(x) = C(k) \exp(ikx) + C(k - G) \exp[i(k - G)x]$ as calculated near the boundary of the first Brillouin zone. One component dominates as we move away from the boundary.

These are the roots for the energy when the wavevector is very close to the zone boundary at $\frac{1}{2}G$. Note the quadratic dependence of the energy on the wavevector \vec{k} . For U negative, the solution $\epsilon(-)$ corresponds to the upper of the two bands, and $\epsilon(+)$ to the lower of the two bands. The two C 's are plotted in Fig. 10.

NUMBER OF ORBITALS IN A BAND

Consider a linear crystal constructed of an even number N of primitive cells of lattice constant a . In order to count states we apply periodic boundary conditions to the wavefunctions over the length of the crystal. The allowed values of the electron wavevector k in the first Brillouin zone are given by (2):

$$k = 0 ; \pm \frac{2\pi}{L} ; \pm \frac{4\pi}{L} ; \dots ; \frac{N\pi}{L} . \quad (53)$$

We cut the series off at $N\pi/L = \pi/a$, for this is the zone boundary. The point $-N\pi/L = -\pi/a$ is not to be counted as an independent point because it is connected by a reciprocal lattice vector with π/a . The total number of points is exactly N , the number of primitive cells.

Each primitive cell contributes exactly one independent value of k to each energy band. This result carries over into three dimensions. With account taken of the two independent orientations of the electron spin, there

are $2N$ independent orbitals in each energy band. If there is a single atom of valence, one in each primitive cell, the band can be half filled with electrons. If each atom contributes two valence electrons to the band, the band can be exactly filled. If there are two atoms of valence, one in each primitive cell, the band can also be exactly filled.

Metals and Insulators

If the valence electrons exactly fill one or more bands, leaving others empty, the crystal will be an insulator. An external electric field will not cause current flow in an insulator. (We suppose that the electric field is not strong enough to disrupt the electronic structure.) Provided that a filled band is separated by an energy gap from the next higher band, there is no continuous way to change the total momentum of the electrons if every accessible state is filled. Nothing changes when the field is applied. This is quite unlike the situation for free electrons for which \mathbf{k} increases uniformly in a field (Chapter 6).

A crystal can be an insulator only if the number of valence electrons in a primitive cell of the crystal is an even integer. (An exception must be made for electrons in tightly bound inner shells which cannot be treated by band theory.) If a crystal has an even number of valence electrons per primitive cell, it is necessary to consider whether or not the bands overlap in energy. If the bands overlap in energy, then instead of one filled band giving an insulator, we can have two partly filled bands giving a metal (Fig. 11).

The alkali metals and the noble metals have one valence electron per primitive cell, so that they have to be metals. The alkaline earth metals have two valence electrons per primitive cell; they could be insulators, but the bands overlap in energy to give metals, but not very good metals. Diamond, silicon, and germanium each have two atoms of valence four, so that there are

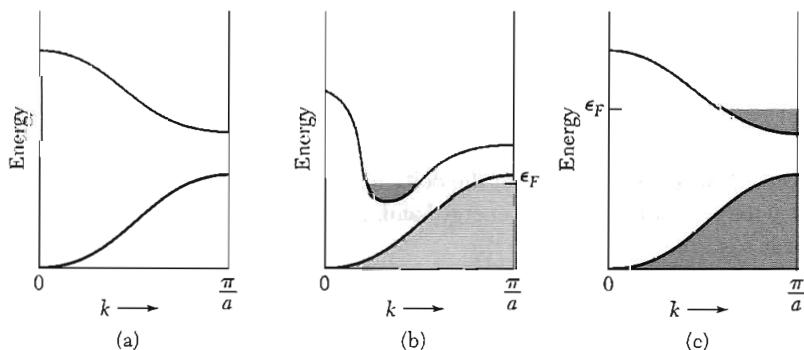


Figure 11 Occupied states and band structures giving (a) an insulator, (b) a metal or a semimetal because of band overlap, and (c) a metal because of electron concentration. In (b) the overlap need not occur along the same directions in the Brillouin zone. If the overlap is small, with relatively few states involved, we speak of a semimetal.

eight valence electrons per primitive cell; the bands do not overlap, and the pure crystals are insulators at absolute zero.

SUMMARY

- The solutions of the wave equation in a periodic lattice are of the Bloch form $\psi_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_{\mathbf{k}}(\mathbf{r})$, where $u_{\mathbf{k}}(\mathbf{r})$ is invariant under a crystal lattice translation.
- There are regions of energy for which no Bloch function solutions of the wave equation exist (see Problem 5). These energies form forbidden regions in which the wavefunctions are damped in space and the values of the \mathbf{k} 's are complex, as pictured in Fig. 12. The existence of forbidden regions of energy is prerequisite to the existence of insulators.
- Energy bands may often be approximated by one or two plane waves: for example, $\psi_k(x) \approx C(k)e^{ikx} + C(k - G)e^{i(k-G)x}$ near the zone boundary at $\frac{1}{2}\mathbf{G}$.
- The number of orbitals in a band is $2N$, where N is the number of primitive cells in the specimen.

Problems

- 1. Square lattice, free electron energies.** (a) Show for a simple square lattice (two dimensions) that the kinetic energy of a free electron at a corner of the first zone is higher than that of an electron at midpoint of a side face of the zone by a factor of 2. (b) What is the corresponding factor for a simple cubic lattice (three dimensions)? (c) What bearing might the result of (b) have on the conductivity of divalent metals?
- 2. Free electron energies in reduced zone.** Consider the free electron energy bands of an fcc crystal lattice in the approximation of an empty lattice, but in the reduced zone scheme in which all \mathbf{k} 's are transformed to lie in the first Brillouin zone. Plot roughly in the [111] direction the energies of all bands up to six times the lowest band energy at the zone boundary at $\mathbf{k} = (2\pi/a)(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$. Let this be the unit of energy. This problem shows why band edges need not necessarily be at the zone center. Several of the degeneracies (band crossings) will be removed when account is taken of the crystal potential.
- 3. Kronig-Penney model.** (a) For the delta-function potential and with $P \ll 1$, find at $k = 0$ the energy of the lowest energy band. (b) For the same problem find the band gap at $k = \pi/a$.
- 4. Potential energy in the diamond structure.** (a) Show that for the diamond structure the Fourier component U_G of the crystal potential seen by an electron is equal to zero for $\mathbf{G} = 2\mathbf{A}$, where \mathbf{A} is a basis vector in the reciprocal lattice referred to the conventional cubic cell. (b) Show that in the usual first-order approximation to the solutions of the wave equation in a periodic lattice the energy gap vanishes at the zone boundary plane normal to the end of the vector \mathbf{A} .

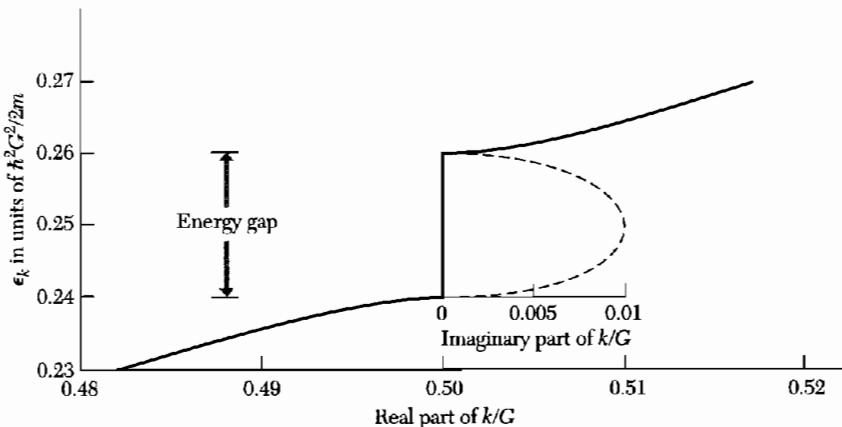


Figure 12 In the energy gap there exist solutions of the wave equation for complex values of the wavevector. At the boundary of the first zone the real part of the wavevector is $\frac{1}{2}G$. The imaginary part of k in the gap is plotted in the approximation of two plane waves, for $U = 0.01 \hbar^2 G^2 / 2m$. In an infinite unbounded crystal the wavevector must be real, or else the amplitude will increase without limit. But on a surface or at a junction there can exist solutions with complex wavevector.

***5. Complex wavevectors in the energy gap.** Find an expression for the imaginary part of the wavevector in the energy gap at the boundary of the first Brillouin zone, in the approximation that led to Eq. (46). Give the result for the $\text{Im}(k)$ at the center of the energy gap. The result for small $\text{Im}(k)$ is

$$(\hbar^2/2m)[\text{Im}(k)]^2 \approx 2mU^2/\hbar^2 G^2 .$$

The form as plotted in Fig. 12 is of importance in the theory of Zener tunneling from one band to another in the presence of a strong electric field.

6. Square lattice. Consider a square lattice in two dimensions with the crystal potential

$$U(x,y) = -4U \cos(2\pi x/a) \cos(2\pi y/a) .$$

Apply the central equation to find approximately the energy gap at the corner point $(\pi/a, \pi/a)$ of the Brillouin zone. It will suffice to solve a 2×2 determinantal equation.

*This problem is somewhat difficult.