6 The WKB Approximation. Electronic Properties of Solids

In real situations, the potentials are usually more complicated than those studied in the previous chapter. However, exact solutions of the Schrödinger equation for such potentials do not generally exist, so it is necessary to develop specific solution methods. Depending on the nature of the problem and the purpose of the study, it is usually convenient to use an approximation method that provides a sufficiently simple and satisfactory solution. Several approximation methods exist, and in this book we will have the opportunity to draw on some of them.

In the first part of this chapter we present an approximation technique developed simultaneously and independently in 1926 by G. Wentzel (1896–1978), H. A. Kramers (1894–1952), and L. Brillouin (1889–1969), from whose surnames the acronym WKB is derived. The name WKB refers specifically to the method when applied to QM. More generally, it is a procedure for constructing approximate solutions to linear differential equations with spatially varying coefficients, known in mathematics as the Liouville–Green[1,2] method. In QM, the WKB method is extremely useful for 1D problems, although it is occasionally used for problems in more than one dimension. It is particularly suitable when the particle is in a sufficiently energetic state that its behavior can be considered to be close to classical almost everywhere, in the sense explained in what follows. For this reason, it is called a *semiclassical approximation* method. The meaning of the term semiclassical in this context should be clearly distinguished from its meaning in the context of *semiclassical theory*, where quantized matter is studied in interaction with classical (nonquantized) radiation.

The rest of the chapter is devoted to a discussion of the basic electronic properties of solids and some important applications of these properties, which are easily explained using the results obtained in the first part.

6.1 The Semiclassical Approach

To solve the stationary Schrödinger equation for the wave function ψ

$$-\frac{\hbar^2}{2m}\nabla^2\psi + V\psi = E\psi \tag{6.1}$$

[1] Joseph Liouville (1809–1882) was a notable French mathematician who made important contributions to several branches of mathematics.

[2] George Green (1793–1841) was an outstanding British mathematical physicist who, among several other contributions, introduced the concept of Green's functions and applied mathematics to the study of electricity and magnetism.

https://doi.org/10.1017/9781009679633.007 Published online by Cambridge University Press

with the WKB method, we write ψ in the form[3]

$$\psi = e^{iS/\hbar}, \tag{6.2}$$

where $S(\boldsymbol{x})$ is a function of $\boldsymbol{x}$ with dimensions of action, usually complex. To determine S we substitute (6.2) into (6.1). This gives

$$\nabla^2\psi = \nabla\cdot(\nabla\psi) = \nabla\cdot\left(\frac{i}{\hbar}\psi\nabla S\right) = \frac{i}{\hbar}(\nabla^2 S)\psi - \frac{(\nabla S)^2}{\hbar^2}\psi, \tag{6.3}$$

so Eq. (6.1) becomes

$$\frac{(\nabla S)^2}{2m} - \frac{i\hbar}{2m}\nabla^2 S + V = E. \tag{6.4}$$

Here we introduce the simplifying consideration that the system is in a state such that

$$\hbar|\nabla^2 S| \ll (\nabla S)^2, \tag{6.5}$$

in which case Eq. (6.4) takes the approximate form

$$\frac{(\nabla S)^2}{2m} + V \doteq E, \tag{6.6}$$

where the $\hbar$ has disappeared; this explains the name "semiclassical" for the procedure. Equation (6.6) reminds us of the Hamilton–Jacobi[4,5] equation of classical mechanics for the action $S(\boldsymbol{x})$, whose gradient determines the momentum of the particle

$$\boldsymbol{p} = \nabla S. \tag{6.7}$$

It is sufficient to substitute this last formula into the Hamilton–Jacobi equation to obtain the correct expression for the energy in a (classical) conservative problem. This means that if the condition (6.5) is satisfied, the function S is approximately given by Eq. (6.7).[6] The semiclassical approach starts from the assumption that under appropriate conditions (basically, that condition (6.5) is met) the quantum problem can be solved approximately by finding quantum corrections to the classical action and using this corrected action to determine the wave function.

Since the corrections we are looking for are caused by the term containing $\hbar\nabla^2 S$, which is proportional to $\hbar$, they in turn must depend on $\hbar$. Based on this, we propose that the solution of Eq. (6.4) can be expressed as an expansion of S in powers of $\hbar$,[7]

$$S = S_0 + i\hbar S_1 + (i\hbar)^2 S_2 + \ldots, \tag{6.8}$$

[3] The solution of the stationary wave equation is often written in optics in the form $\exp iS(\boldsymbol{x})$, where S is known as the *eikona/*, from the Greek word for icon (image). By extension, the WKB method is sometimes referred to as an expansion in terms of the eikonal, that is, a term corresponding to a wave, either stationary or propagating.

[4] Sir William Rowan Hamilton (1805–1865) was an eminent British mathematician, physicist, and astronomer, who made many contributions, including Hamiltonian mechanics.

[5] Carl Gustav Jacob Jacobi (1804–1851) was an eminent German mathematician; it would take a long list to recall all the achievements that bear his name, including a crater on the Moon.

[6] A word of caution, however: while the Hamilton–Jacobi equation of classical mechanics describes the deterministic motion of a single particle, Eq. (6.4) and its approximate version (6.6) refer to an ensemble of particles. The analogy is therefore strictly formal.

[7] This assumes that the solution can be expressed by an analytical function of Planck's constant, which is not always the case.

https://doi.org/10.1017/9781009679633.007 Published online by Cambridge University Press

where none of the functions S_j, $j = 0, 1, 2, \ldots$ contains $\hbar$. Substitution of the gradient of this expression into (6.4) gives

$$\frac{1}{2m}[(\nabla S_0)^2 + 2i\hbar\nabla S_0 \cdot \nabla S_1 - 2\hbar^2\nabla S_0 \cdot \nabla S_2 - \hbar^2(\nabla S_1)^2 + \ldots] - \frac{i\hbar}{2m}[\nabla^2 S_0 + i\hbar\nabla^2 S_1 + \ldots] + V - E = 0. \tag{6.9}$$

By canceling the coefficients of each power of $\hbar$ separately, we obtain a hierarchy of differential equations for the S_j, $j = 0, 1, 2, \ldots$,

$$\frac{(\nabla S_0)^2}{2m} + V = E, \tag{6.10}$$

$$\nabla S_0 \cdot \nabla S_1 - \tfrac{1}{2}\nabla^2 S_0 = 0, \tag{6.11}$$

$$2\nabla S_0 \cdot \nabla S_2 + (\nabla S_1)^2 - \nabla^2 S_1 = 0, \tag{6.12}$$

and so on. The first of these equations coincides with the (classical) Hamilton–Jacobi equation, therefore we take for S_0 the classical action of the system, which is obtained from an integration of the classical momentum,

$$\nabla S_0 = \boldsymbol{p}. \tag{6.13}$$

Once S_0 is determined, Eq. (6.11) should allow us to obtain S_1; substituting S_0 and S_1 in Eq. (6.12), we can determine S_2 in principle, and so on. Normally it is sufficient to know S to the first order in $\hbar$, which is achieved by solving only the first two equations of the hierarchy.

In two- or three-dimensional problems, additional considerations are required to solve Eqs. (6.10)–(6.12), hence the limited applicability of the method in such cases. In the following we restrict ourselves to the one-dimensional case, where Eqs. (6.10) and (6.11) take the simple form

$$\left(S_0'\right)^2 = 2m(E - V) = p^2, \tag{6.14}$$

$$S_1' = \frac{S_0''}{2S_0'} = \tfrac{1}{2}\frac{d}{dx}\ln S_0' = \frac{d}{dx}\ln\sqrt{p}. \tag{6.15}$$

Integrating, we get

$$S_0 = \pm\int^x p\,dx; \quad S_1 = \ln\sqrt{p} = -\ln\frac{1}{\sqrt{p}}. \tag{6.16}$$

Therefore, up to linear terms in $\hbar$ (classical solution plus first quantum correction) and to a normalization factor, the wave function has the general form

$$\psi = e^{(i/\hbar)(S_0 + i\hbar S_1)} = e^{iS_0/\hbar - S_1} = \exp\left[\pm(i/\hbar)\int^x p\,dx + \ln\frac{1}{\sqrt{p}}\right], \tag{6.17}$$

https://doi.org/10.1017/9781009679633.007 Published online by Cambridge University Press

that is,[8]

$$\psi = \frac{1}{\sqrt{|p|}} \exp\left[\pm i/\hbar \int^x p\, dx\right]. \tag{6.18}$$

To establish the range of validity of this approximation, we note that Eqs. (6.10) and (6.13) allow us to write the magnitude of p in its classical form,

$$p = \sqrt{2m(E - V)}, \tag{6.19}$$

and from (6.5) and (6.13) it follows that the semiclassical method is applicable only if

$$\hbar \left|\frac{dp}{dx}\right| \ll p^2. \tag{6.20}$$

From Eq. (6.19) we get

$$\frac{dp}{dx} = -\sqrt{\frac{m}{2(E - V)}}\frac{dV}{dx} = \frac{mF}{p}, \tag{6.21}$$

where $F(x)$ represents the external force acting on the particle; substituting in (6.20), we get

$$|\hbar m F/p^3| \ll 1. \tag{6.22}$$

This condition is only fulfilled in regions of x where the force F is "small" and the classical momentum p is "large." From this we conclude that quantum effects are particularly noticeable: (a) where the potential V varies strongly; (b) in regions close to the classical turning points, for which $V = E$ and classical particles stop, since they cannot move to regions where $V > E$. In such regions, the semiclassical formulae are not valid.

Figure 6.1 shows a smoothly varying potential, to which we can apply the WKB method. The shaded zones around the turning points x_1 and x_2 are to be excluded according to the preceding discussion. To determine the width of the exclusion zones we proceed as follows. Near the turning point x_1 we can approximate the potential with a first-order Taylor expansion:

$$V(x) = V(x_1) + (x - x_1)V' = E - (x - x_1)F, \tag{6.23}$$

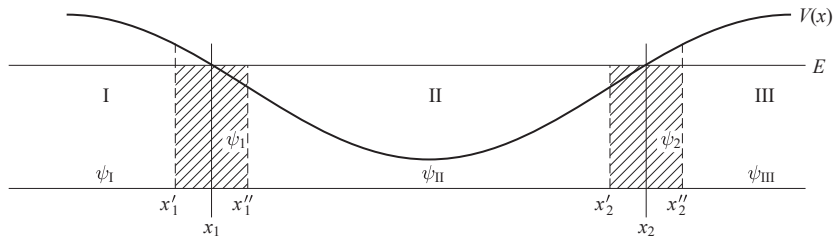


Figure 6.1 In regions I, II, and III the WKB functions are applicable; the shaded areas represent the exclusion zones.

[8] Note that in the present approximation the "classical action" $\int^x p\, dx$ is everywhere measured with respect to Planck's constant, a most fundamental reminder of the role of $\hbar$. Some authors refer to the semiclassical limit as the limit $\hbar \to 0$; however, $\hbar$ is a universal constant with a fixed value. Expressions that condition the validity of the approximation, such as (6.20) or (6.22), must be read as constraints on the dynamics, not on $\hbar$.

https://doi.org/10.1017/9781009679633.007 Published online by Cambridge University Press

since at $x = x_1$, $V = E$. Substituting this expression in Eq. (6.19), we obtain for $x > x_1$ (to the right of the turning point),

$$p = \sqrt{2mF(x - x_1)}, \tag{6.24}$$

and substituting this result in (6.22), we obtain

$$|x - x_1| \gg \tfrac{1}{2} \left| \frac{\hbar^2}{mF} \right|^{1/3}, \tag{6.25}$$

which means that the semiclassical solution applies outside this zone, $x > x_1''$ in Fig. 6.1. Using this criterion we eliminate from our analysis the neighboring zones of x_1 and x_2, now the WKB approach can be safely applied in the regions marked I, II, and III.

The solution (6.18) applies to the case $E > V$, since p was taken to be real (see Eq. (6.19)); hence it corresponds to the wave function of region II, ψ_{II}. To construct this function, we note that the general solution is the sum of the two imaginary exponential solutions with arbitrary coefficients, which can be rewritten in terms of a trigonometric function using an amplitude and a phase as integration constants, with the origin for the latter at x_1,

$$\psi_{II}(x) = \frac{A_2}{\sqrt{p}} \sin\left[\frac{1}{\hbar}\int_{x_1}^{x} p(x')\,dx' + \alpha_2\right], \quad x_1'' \le x \le x_2'. \tag{6.26}$$

The factor $\sqrt{p}$ in the denominator of ψ_{II} means that the fraction of particles $\rho dx \sim \left(|A_2|^2/p\right) dx$ that are on average in the interval dx becomes smaller as the flux velocity becomes larger. This is intuitively clear from the classical perspective, where the average time a particle spends in the interval $(x, x+dx)$ is inversely proportional to the average velocity at x.

In regions I and III, $V > E$ and p is imaginary, so the exponentials are real; therefore we must write for $x \ge x_2''$

$$\psi_{III} = \frac{1}{\sqrt{|p|}}\left[A_3 \exp\left(-\frac{1}{\hbar}\int_{x_2}^{x} |p(x')|\,dx'\right) + B_3 \exp\left(\frac{1}{\hbar}\int_{x_2}^{x} |p(x')|\,dx'\right)\right] \tag{6.27}$$

and an analogous expression in region I for $x \le x_1'$. In these expressions $|p|$ is

$$|p| = \sqrt{2m(V - E)}. \tag{6.28}$$

These equations show that in the classically allowed regions, the WKB wave function is oscillatory, while it tends to decay exponentially in the classically forbidden regions.

We now need to stitch the piecewise solutions together to ensure that they are part of the same wave function. The problem is not so simple, because the preceding solutions are not valid at the turning points. But just near the turning points the potential can be approximated with a first-order Taylor expansion, like (6.23), which, when introduced into the Schrödinger equation, gives a solution in terms of Bessel functions. Then, for example, the solution ψ_1 valid in the exclusion zone $x_1' < x < x_1''$, is tied to ψ_I at x_1' and to ψ_{II} at x_1'', which establishes a relation between the respective coefficients A_i, B_i. Similarly, the solution ψ_2 valid for $x_2' < x < x_2''$, is tied to ψ_{II} at x_2' and to ψ_{III} at x_2'' (see Fig. 6.1). Since this is a purely technical subject that can be consulted in the literature, we will omit it here and only present the final results. The pair of solutions

https://doi.org/10.1017/9781009679633.007 Published online by Cambridge University Press

tied to both sides of x_2, obtained under the hypothesis that in region III the solution is exponentially decreasing, is

$$\psi_{II}(x) = \frac{A}{\sqrt{p}} \sin\left[\frac{1}{\hbar}\int_x^{x_2} p(x')\,dx' + \frac{\pi}{4}\right], \quad x < x_2', \tag{6.29}$$

$$\psi_{III}(x) = \frac{A}{2\sqrt{|p|}} \exp\left[-\frac{1}{\hbar}\int_{x_2}^{x} |p(x')|\,dx'\right], \quad x > x_2''. \tag{6.30}$$

A similar result is obtained for the solutions around the turning point x_1, but with the integration limits reversed, so that the lower limit is always less than the upper one.

Exercise E6.1. Quantization in a potential well

Apply the WKB formulae to a particle bound by an arbitrary but smoothly varying potential well, similar to that shown in Fig. 6.1, to derive the action quantization condition and an (approximate) expression for the normalized wave function inside the well.

Solution. We can write the WKB wave function inside the well in two equivalent forms,

$$\psi = \frac{A'}{\sqrt{p}} \sin\left[\frac{1}{\hbar}\int_{x_1}^{x} p\,dx + \frac{\pi}{4}\right], \tag{6.31}$$

$$\psi = \frac{A''}{\sqrt{p}} \sin\left[\frac{1}{\hbar}\int_{x}^{x_2} p\,dx + \frac{\pi}{4}\right]. \tag{6.32}$$

From the second expression it follows that

$$\begin{aligned}\psi &= \frac{A''}{\sqrt{p}} \sin\left[\frac{1}{\hbar}\left(\int_{x_1}^{x_2} - \int_{x_1}^{x}\right) p\,dx + \frac{\pi}{2} - \frac{\pi}{4}\right] \\ &= -\frac{A''}{\sqrt{p}} \sin\left[\frac{1}{\hbar}\int_{x_1}^{x} p\,dx + \frac{\pi}{4} - \frac{1}{\hbar}\int_{x_1}^{x_2} p\,dx - \frac{\pi}{2}\right],\end{aligned} \tag{6.33}$$

and for this to coincide with Eq. (6.31) it must hold that

$$\frac{1}{\hbar}\int_{x_1}^{x_2} p\,dx + \frac{\pi}{2} = \pi(n+1), \quad n = 0, 1, 2, \ldots \tag{6.34}$$

$$A' = (-)^n A''. \tag{6.35}$$

In writing (6.34) we took into account that $\int_{x_1}^{x_2} p\,dx > 0$, so the right-hand side cannot be zero. The quantization condition (6.34) shows that bound particles have a discrete spectrum regardless of the shape of the well. This is a general rule, of which we have found examples in Chapter 5 and will find others later, an important one being that of the atom. Since p has a different sign on the forward and the return paths, we can write the quantization rule as a closed integral over a complete period,

$$\oint p\,dx = 2\pi\hbar\left(n + \tfrac{1}{2}\right). \tag{6.36}$$

https://doi.org/10.1017/9781009679633.007 Published online by Cambridge University Press

This result (which is exact up to terms of order $\hbar$) agrees with the Wilson–Sommerfeld rule (see Section 2.6), except for the presence of the additional term $\frac{1}{2}h$, which is reminiscent of the zero-point energy.[9]

Equations (6.31) and (6.34) show that the phase of ψ changes from $\frac{1}{4}\pi$ at $x = x_1$ to $\pi(n + \frac{3}{4})$ at $x = x_2$, so the quantum number n determines the number of nodes of the wave function inside the well. Furthermore, by eliminating F from Eqs. (6.24) and (6.25), we obtain

$$|x - x_1| \gg \frac{\hbar}{2p}, \tag{6.37}$$

which confirms that the semiclassical approach applies to states characterized by a large quantum number n.

To determine the normalization constant we consider that since outside the well the wave function decays exponentially, we can neglect its contribution to the normalization integral. We then normalize the ψ given by (6.31) by approximating the square of the sine function, which is a rapidly oscillating function (remember that $n \gg 1$), by its mean value $\frac{1}{2}$, thus

$$\begin{aligned} &|A'|^2 \int_{x_1}^{x_2} \sin^2\left(\frac{1}{\hbar}\int_{x_1}^{x} p(x')\,dx' + \frac{\pi}{4}\right)\frac{dx}{p} \simeq \tfrac{1}{2}|A'|^2 \int_{x_1}^{x_2}\frac{dx}{p} \\ &= \frac{|A'|^2}{2m}\int_{x_1}^{x_2}\frac{dx}{v} = \frac{|A'|^2}{4m}\oint\frac{dx}{v} = \frac{|A'|^2}{4m}T = \frac{|A'|^2}{4m}\frac{2\pi}{\omega} = 1. \end{aligned} \tag{6.38}$$

The period of oscillation $T = \oint dx/v$ has been re-expressed in terms of the frequency $\omega = 2\pi/T$. Taking the normalization constant to be real and positive, we get $A' = \sqrt{2m\omega/\pi}$. The normalized WKB wave function inside the well is thus

$$\psi = \sqrt{\frac{2\omega}{\pi v}}\sin\left(\frac{1}{\hbar}\int_{x_1}^{x} p\,dx + \frac{\pi}{4}\right), \tag{6.39}$$

subject to the quantization condition (6.36).

Exercise E6.2. Quantization in a multidimensional well

Obtain the quantization conditions for particles in a multidimensional potential well using the WKB method.[10]

Solution. We initially write the WKB solution in the form $\psi = Ae^{iS}$. For this ψ to be satisfactory, we must require that A be single-valued; the action S in turn may be multivalued, provided that the difference between two of its values satisfies the condition

$$\Delta S \equiv \oint_C \nabla S\,ds = 2\pi n, \quad n = 1, 2, 3, \ldots \tag{6.40}$$

[9] When applied to central potentials, the results of the WKB approximation are often improved by the replacement $l(l+1) \to (l+\frac{1}{2})^2$, where l represents the angular momentum.

[10] The WKB problem in more than one dimension, if not separable, is considerably more complicated than the 1D case and involves the topological properties of phase space. Due to the complications of the full solution, this Exercise is given an informal treatment.

https://doi.org/10.1017/9781009679633.007 Published online by Cambridge University Press

for any closed contour C of integration in x-space. However, it often happens that the function A is not single-valued on certain surfaces. In these cases we can proceed as follows. Let A and $A + \Delta A$ be two possible values of the function A at a given point. We rewrite the wave function in the form

$$\psi = (A + \Delta A)e^{iS} = Ae^{i(S - i\Delta \ln A)}, \tag{6.41}$$

so that we have translated the change in A into a change in phase. Condition (6.40) is now rewritten in the form

$$\int_C \nabla S\, ds = 2\pi(n + \frac{i}{2\pi}\Delta \ln A). \tag{6.42}$$

Considering the general results of the WKB method, which show that $\int_C \nabla S ds = (1/\hbar)\oint_C p_i dx_i$, we see that the quantization rules have the general form

$$\oint p_i\, dx_i = 2\pi\hbar(n + \frac{i}{2\pi}\Delta \ln A). \tag{6.43}$$

There are as many quantization integrals as there are nonreducible contours in the problem (two contours are reducible if it is possible to continuously transform one into the other without crossing any singularity of the integrand). Contours that are reducible to a point can be eliminated because the integral over them is zero.

As an example of Eq. (6.43) we consider the case where A changes sign, as happens in the case of a perfectly reflecting wall. Since

$$\frac{i}{2\pi}\Delta \ln A = \frac{i}{2\pi}(\ln(-A) - \ln A) = \frac{i}{2\pi}\ln(-1) = \frac{i}{2\pi}\ln e^{-i\pi} = \tfrac{1}{2}, \tag{6.44}$$

the quantization rule in this case contains a half-integer.

6.2 Transmission Across a Barrier. Alpha Decay

Consider a smoothly varying repulsive potential, as illustrated in Fig. 6.2. We shall use the WKB method to study the behavior of a beam of particles approaching the barrier from the left. Our first objective is to determine the reflection and transmission coefficients of particles with energy E less than $V_{\max}$. The following presentation is rather condensed; it is recommended that the student fill in the omitted steps.

To solve this problem, we need the connection formulae between the solutions in regions I and II and II and III, respectively. These formulae are obtained by following the same procedure used to establish the pair of equations (6.29)–(6.30). The result for regions I and II is

$$\psi_I = \frac{1}{\sqrt{p}}\left[A_1 e^{i\left(\frac{1}{\hbar}\int_x^{x_1} p\,dx + \frac{\pi}{4}\right)} + B_1 e^{-i\left(\frac{1}{\hbar}\int_x^{x_1} p\,dx + \frac{\pi}{4}\right)}\right], \quad x < x_1 \tag{6.45}$$

$$\psi_{II} = \frac{1}{\sqrt{|p|}}\left[(A_1 + B_1)e^{\frac{1}{\hbar}\int_{x_1}^{x} |p|dx} + \frac{i}{2}(A_1 - B_1)e^{-\frac{1}{\hbar}\int_{x_1}^{x} |p|dx}\right], \quad x > x_1 \tag{6.46}$$

https://doi.org/10.1017/9781009679633.007 Published online by Cambridge University Press

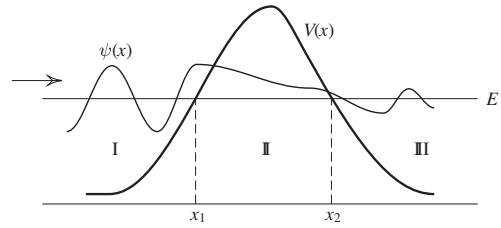


Figure 6.2 A smoothly varying potential barrier. The general form of the solution is shown for $E < V_{max}$.

Since in region III the particles travel only to the right, for the second pair of regions we get

$$\psi_{III} = \frac{1}{\sqrt{p}} A_3 e^{i\left(\frac{1}{\hbar}\int_{x_2}^{x} p dx + \frac{\pi}{4}\right)}, \quad x > x_2 \tag{6.47}$$

$$\psi_{II} = \frac{A_3}{\sqrt{|p|}}\left[e^{\frac{1}{\hbar}\int_{x}^{x_2}|p|dx} - \tfrac{1}{2} i e^{-\frac{1}{\hbar}\int_{x}^{x_2}|p|dx}\right]. \quad x < x_2 \tag{6.48}$$

With the integral in the exponents written as

$$\int_{x}^{x_2} |p|dx = \int_{x_1}^{x_2} |p|dx - \int_{x_1}^{x} |p|dx \tag{6.49}$$

and the definition as

$$J \equiv \frac{1}{\hbar}\int_{x_1}^{x_2} |p|dx = \frac{\sqrt{2m}}{\hbar}\int_{x_1}^{x_2} \sqrt{V-E}dx, \tag{6.50}$$

ψ_{II} can be rewritten as

$$\psi_{II} = \frac{A_3}{\sqrt{|p|}}\left[e^{J-\frac{1}{\hbar}\int_{x_1}^{x}|p|dx} - \tfrac{1}{2} i e^{-J+\frac{1}{\hbar}\int_{x_1}^{x}|p|dx}\right]. \tag{6.51}$$

Comparing the coefficients in this expression with those in (6.46), we see that

$$A_1 + B_1 = -\tfrac{1}{2} i A_3 e^{-J}; \quad \tfrac{1}{2} i(A_1 - B_1) = A_3 e^{J}, \tag{6.52}$$

from which it follows that

$$A_1 = -i A_3(e^{J} + \tfrac{1}{4} e^{-J}), \quad B_1 = i A_3(e^{J} - \tfrac{1}{4} e^{-J}). \tag{6.53}$$

Using the definitions for the transmission and reflection coefficients introduced in Chapter 5, we obtain

$$T = \frac{|j_{\text{trans}}|}{|j_{\text{inc}}|} = \frac{|A_3|^2}{|A_1|^2} = (e^{J} + \tfrac{1}{4} e^{-J})^{-2}, \tag{6.54}$$

$$R = \frac{|j_{\text{ref}}|}{|j_{\text{inc}}|} = \frac{|B_1|^2}{|A_1|^2} = \left(\frac{e^{J} - \frac{1}{4} e^{-J}}{e^{J} + \frac{1}{4} e^{-J}}\right)^2. \tag{6.55}$$

Usually $J \gg 1$ since the WKB approximation is satisfactory only for wide barriers, so we can approximate the transmission coefficient as

$$T \approx e^{-2J} = e^{-\frac{2}{\hbar}\int_{x_1}^{x_2}\sqrt{2m(V-E)}dx}. \tag{6.56}$$

https://doi.org/10.1017/9781009679633.007 Published online by Cambridge University Press

This is a very useful formula that generalizes the expression of T for arbitrary but smooth potentials. It shows that the transparency of a barrier to a particle depends very strongly on its energy and its mass; for example, if electrons are replaced by protons of the same energy, the transmission coefficient decreases by a factor of the order of $e^{\sqrt{1840}} \sim e^{42} \sim 10^{18}$. We will see some examples in what follows that illustrate the importance of this phenomenon.

Note that Fig. 6.2 again shows the existence of a nonlocal feature: the incident wave function is distorted to adapt to the wave function in region II. This adaptation involves a small zone *outside* the potential, as if to announce to the incoming wave that a change in the potential function is imminent. We must bear in mind that, as discussed in Section 5.4, the wave function provides a statistical, time-independent description, when the ensemble has had sufficient occasion to reach a stationary distribution.

6.2.1 Nuclear alpha decay

Some atomic nuclei decay spontaneously (without any external intervention) by emitting α particles, that is, ^{4}He nuclei, with two protons and two neutrons. The energy with which the α particles are emitted from the radioactive nucleus is characteristic of each isotope and is of the order of a few MeV, that is, much lower than the Coulomb repulsion energy between the α particle and the $Z-2$ charged daughter nucleus (Z is the charge of the parent nucleus). This led E. Condon,[11] Gurney,[12] and, independently, Gamow,[13] in 1928 to understand α decay as a phenomenon of particle transmission through a potential barrier. The theory based on this hypothesis was not only the first physical application of quantum tunneling and one of the first verifications of Schrödinger's theory, but also showed that, with care, QM could be extended even to the scale of atomic nuclei.

To study nuclear alpha decay, we model the potential as shown in Fig. 6.3. R is the radius of the daughter nucleus (of charge $Z_2 = Z-2$); for $r < R$ the particle is attracted by the nucleus with a potential assumed to be rectangular for simplicity, and for $r > R$ there is the Coulomb repulsion between the α particle (of charge $Z_1 = 2$) and the daughter nucleus. Assuming that the particle is in its ground state with energy E, the transmission coefficient is

$$T = \exp\left[-\frac{2\sqrt{2M}}{\hbar}\int_R^{R1}\sqrt{V-E}\,dr\right], \tag{6.57}$$

[11] Edward Uhler Condon (1902–1974) was a US nuclear physicist and pioneer of QM. During the McCarthy era, when efforts were made to root out communist sympathizers in the United States, Condon was targeted by the House Un-American Activities Committee for being a "follower" of a "new revolutionary movement," namely QM. Condon defended himself with a notorious commitment to physics and science. A lunar crater is named after him.

[12] Ronald Wilfrid Gurney (1898–1953) was a British theoretical physicist who initiated the study of radioactive decay.

[13] Gueorgiy Antonovich Gamow (1904–1968) was a prominent Soviet–US polymath, theoretical physicist, and cosmologist, author of a large collection of popular science books, including *One Two Three … Infinity* and the *Mr Tompkins* series, still in print after more than fifty years. Gamow made important scientific contributions, such as the liquid-drop model of the atomic nucleus.

https://doi.org/10.1017/9781009679633.007 Published online by Cambridge University Press

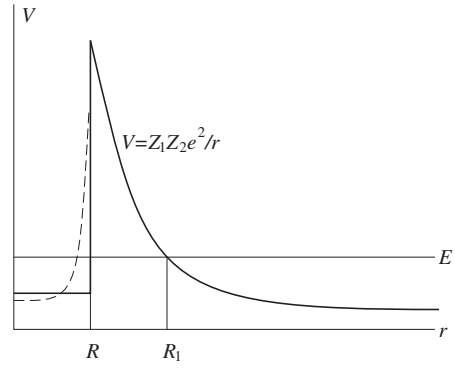


Figure 6.3 Idealized potential barrier for alpha decay. The dotted line inside the nucleus shows a more realistic representation of the barrier.

where M is the mass of the α particle and V is given by

$$V = \frac{Z_1 Z_2 e^2}{r}, \quad R < r < R_1. \tag{6.58}$$

The turning point at R_1 is given by the condition $V(R_1) = Z_1 Z_2 e^2/R_1 = E$. The integral can be easily calculated with the change of variable $r/R_1 = \sin^2\varphi$,

$$\int_R^{R_1} \sqrt{\frac{Z_1 Z_2 e^2}{r} - E}\, dr = \sqrt{E} \int_R^{R_1} \sqrt{\frac{R_1}{r} - 1}\, dr = \tfrac{1}{2} R_1 \sqrt{E}(\pi - 2\varphi_0 - \sin 2\varphi_0), \tag{6.59}$$

where $R/R_1 = \sin^2\varphi_0$. By making the approximation $\sin\varphi_0 \approx \varphi_0$, which is valid for wide potential barriers, we obtain

$$\int_R^{R_1} \sqrt{V - E}\, dr = R_1 \sqrt{E} \left[\frac{\pi}{2} - 2\sqrt{\frac{R}{R_1}} \right]. \tag{6.60}$$

Substituting this result in Eq. (6.57), we get with $Z_1 = 2$, $\ln T = -2\pi e^2 Z_2 \sqrt{2M/\hbar^2 E} + 8e\sqrt{Z_2 M R/\hbar^2}$, which can be written as

$$\ln T = c - \frac{a}{\sqrt{E}}, \tag{6.61}$$

with the parameters a and c given by

$$a = \frac{2\pi e^2}{\hbar}\sqrt{2M} Z_2, \; c = \frac{8e}{\hbar}\sqrt{Z_2 M R}. \tag{6.62}$$

An estimate of the decay rate λ is obtained by considering that, since the particle is confined in a "box" of width R and has the lowest possible energy, only half of a de Broglie wavelength is accommodated within the box, giving $R \sim h/p$ or, equivalently, $v \sim h/MR$, from which

$$\lambda \simeq \frac{\hbar}{MR^2} T. \tag{6.63}$$

Taking the logarithm and combining with the previous results, we get

$$\ln \frac{1}{\lambda} = -b + \frac{a}{\sqrt{E}}, \tag{6.64}$$

https://doi.org/10.1017/9781009679633.007 Published online by Cambridge University Press

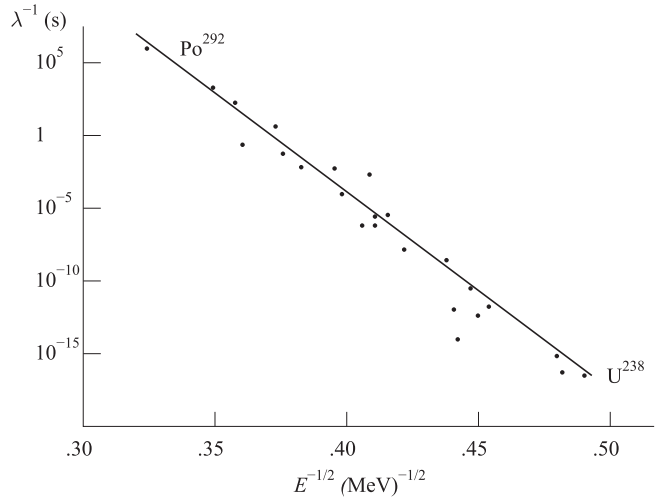


Figure 6.4 Decay constant for a wide range of radioactive isotopes as a function of alpha particle emission energy.

with b given by

$$b = \ln \frac{\hbar}{MR^2} + \frac{8e}{\hbar}\sqrt{Z_2 MR}. \tag{6.65}$$

The analysis of the parameters a and b calculated for different nuclei shows that b varies very slowly with the atomic number Z_2 (in a first approximation it can be taken as a constant), while a varies practically linearly with Z_2, as in Eq. (6.62). In this way, it is verified that, despite the many simplifications made, the theory is able to give a satisfactory qualitative and quantitative explanation of the phenomenon of nuclear α decay.

Note that according to Eq. (6.64), the mean life varies strongly with energy, due to the exponential dependence. For example, in Fig. 6.4 we see that as $1/\sqrt{E}$ goes from about 0.34 (E in MeV) to about 0.5 (less than twice), it changes by 23 *orders of magnitude*. This explains why some radioactive elements are much less stable than others, rendering them very useful for radiometric dating of a wide variety of materials of biological, geological or archaeological interest. For example, C^{14} has a mean life of roughly 8,000 years, while its other isotopes are stable; this makes it very suitable for dating materials (biological or mineral) with a high natural carbon content, since their C^{14}/C^{12} ratio decreases clearly with time.

6.2.2 Time delay by a potential barrier

In the first part of this section we studied the behavior of low-energy particles incident on a repulsive barrier for $E < V_{\max}$, and we saw that most of them are reflected by the potential, although some manage to pass through the barrier. If we increase the incident energy E, from the moment it is equal to some maximum potential, resonance phenomena similar to those studied in the previous chapter begin to appear. The main difference is that quantities such as $\sin qx$ are now replaced by $\sin \int^x q(x)dx$, and so on. This is easy to understand if we think of the potential $V(x)$

https://doi.org/10.1017/9781009679633.007 Published online by Cambridge University Press

as a sequence of a large number of very narrow rectangular barriers placed side by side (see Problem P6.1).

If the incident energy is increased even more, so that $E \gg V_{\max}$, almost all the particles pass over the barrier. Since in this case the scattering is small, we can assume that a high-energy incident packet is transmitted almost without distortion. Under these conditions, the effect of the potential will essentially manifest itself in a delay of the packet. We will show that the delay predicted by quantum theory is given, to a first approximation, by the corresponding classical expression.

This example allows us to study a case of incident packet scattering that deviates from the oversimplified description of free particles described by plane waves. It is interesting to compare this description with that of the example in Section 5.4 in terms of the phase shift of the transmitted wave function, where it was the position of the packet in the asymptotic region that was shifted; here the focus is on the time delay of the arrival of the packet centroid.

Consider a potential that varies smoothly and differs from zero only in a region of width a; from the left and from a very large distance (which is eventually made infinite), particles are launched, forming a narrow packet with mean energy E_0 and initial amplitude $\psi(x,0)$. At a very large distance to the right of the potential (which is eventually considered infinite), there will be a packet of similar shape to the incident one, but with a certain phase that depends on the potential, shifted with respect to the phase it would have in the absence of the potential. To obtain the delay time Δt we first establish its relationship with the associated phase shift δ.

The incoming packet can be expressed in terms of free-particle solutions, which is equivalent to performing a Fourier analysis,

$$\psi(x,t) = \int A(\omega) e^{i(kx-\omega t)}\, dk, \tag{6.66}$$

where ω represents as usual the energy measured in units of $\hbar$,

$$\omega = \frac{E}{\hbar} = \frac{\hbar k^2}{2m}. \tag{6.67}$$

We will consider that the packet consists of particles that are launched from position x_0 with energies distributed around a sharp maximum centered at $E_0 = \hbar\omega_0 = \hbar k_0/2m$,

$$\psi(x,0) = 2\pi\, g(\omega-\omega_0)\delta(x-x_0), \tag{6.68}$$

where the function $g(z)$ is narrow and has a sharp maximum at $z = 0$; the numerical factor 2π is introduced for convenience. On the other hand, from (6.66) with $t = 0$ we obtain

$$\psi(x,0) = \int A(\omega) e^{ikx} dk. \tag{6.69}$$

Taking the inverse transform

$$A(\omega) = \frac{1}{2\pi} \int \psi(x,0) e^{-ikx} dx \tag{6.70}$$

and substituting the value of $\psi(x,0)$ given by (6.68), we get

$$A(\omega) = \int g(\omega-\omega_0)\delta(x-x_0) e^{-ikx} dx = g(\omega-\omega_0) e^{-ikx_0}. \tag{6.71}$$

https://doi.org/10.1017/9781009679633.007 Published online by Cambridge University Press

We introduce this result in (6.66) to obtain

$$\psi_{\text{inc}}(x,t) = \int g(\omega-\omega_0)e^{ik(x-x_0)-i\omega t}dk. \tag{6.72}$$

This is the packet that hits the barrier.

To construct the transmitted packet we do as follows. We have seen in the previous sections and in Chapter 5, that when a plane wave is transmitted by the barrier, its amplitude is modified, but not its frequency. The amplitude of the transmitted wave packet can therefore be written from (6.72) by modifying the amplitude of each component by a factor $S(\omega)$, whose module obviously cannot exceed unity, that is, $|S(\omega)| \leq 1$,

$$\psi_{\text{trans}}(x,t) = \int g(\omega-\omega_0)S(\omega)e^{ik(x-x_0)-i\omega t}dk. \tag{6.73}$$

It is clear that $S(\omega)$ plays the role of a transmission amplitude of the frequency component ω, in the sense that

$$T(\omega) = |S(\omega)|^2. \tag{6.74}$$

Reversing this expression, we can write S in the form

$$S(\omega) = \sqrt{T(\omega)}e^{i\delta(\omega)}. \tag{6.75}$$

This formula says that $\sqrt{T}$ is the factor by which the amplitude of the component of frequency ω is reduced, while $\delta(\omega)$ is the phase shift produced by its delay (if there is no potential, $T = 1$ and $\delta = 0$). We will limit ourselves to the case where the magnitude T varies slowly with the energy, so that we can write approximately $T(\omega) \simeq T(\omega_0)$. We will also assume that the phase can be approximated up to linear terms in the variable ϵ,

$$\epsilon \equiv \omega - \omega_0. \tag{6.76}$$

With these considerations, we write

$$S(\omega) \simeq \sqrt{T(\omega_0)}e^{i\left(\delta(\omega_0)+\epsilon\frac{\partial\delta}{\partial\omega}\right)} = S(\omega_0)e^{i\frac{\partial\delta}{\partial\omega}\epsilon}. \tag{6.77}$$

In turn, as a function of the variable ϵ and up to linear terms, k is written as

$$k = \frac{\sqrt{2mE}}{\hbar} = \sqrt{\frac{2m\omega}{\hbar}} = \sqrt{\frac{2m\omega_0}{\hbar}}\left(1+\frac{\epsilon}{2\omega_0}\right) = k_0 + \frac{m}{p_0}\epsilon. \tag{6.78}$$

We now introduce the approximation (6.77) into (6.73); using (6.76) and (6.78), we expand around k_0 and write

$$\begin{aligned}\psi_{\text{trans}} &= S(\omega_0)\int e^{i\frac{\partial\delta}{\partial\epsilon}\epsilon}g(\epsilon)e^{i(k_0+\frac{m}{p_0}\epsilon)(x-x_0)-it(\omega_0+\epsilon)}\frac{m}{p_0}d\epsilon \\ &= S(\omega_0)e^{ik_0(x-x_0)-i\omega_0 t}\int g(\epsilon)e^{\frac{im}{p_0}\left[x-x_0-\frac{p_0}{m}\left(t-\frac{\partial\delta}{\partial\epsilon}\right)\right]\epsilon}\frac{m}{p_0}d\epsilon.\end{aligned} \tag{6.79}$$

The factor preceding the integral is the central component of the energy packet; we see that its amplitude has indeed been reduced by the factor $S(\omega_0)$. Much more interesting, however, is the integral, since it describes the envelope of the incident packet, but with the time shifted by the quantity

$$\Delta t = \frac{\partial\delta}{\partial\epsilon} = \hbar\frac{\partial\delta}{\partial E}. \tag{6.80}$$

https://doi.org/10.1017/9781009679633.007 Published online by Cambridge University Press

This gives the desired relationship between the phase shift δ and the delay time Δt.

To calculate δ we use the WKB solution for the transmitted wave function

$$\psi_{\text{trans}} = \frac{A}{\sqrt{p}} e^{\frac{i}{\hbar}\int_{x_0}^{x} p dx} \tag{6.81}$$

and rewrite the integral of the exponent as

$$\int_{x_0}^{x} p dx = \int_{x_0}^{x} (p_0 + p - p_0) dx = p_0(x - x_0) + \int_{x_0}^{x} (p - p_0) dx, \tag{6.82}$$

so that the transmitted packet takes the form

$$\psi_{\text{trans}} = B(p_0) e^{i\int_{x_0}^{x} (p-p_0)dx/\hbar}. \tag{6.83}$$

In the absence of the scattering potential, we have $p = p_0$ at every point and the previous expression reduces to $\psi_{\text{trans}} = B(p_0)$; therefore, the factor of $B(p_0)$ represents the additional phase introduced by the potential along the path from x_0 to x. Since the potential is assumed to be localized, we can extend the limits of integration to infinity on both sides to obtain

$$\delta(E) = \int_{-\infty}^{\infty} \frac{p - p_0}{\hbar} dx = \frac{\sqrt{2m}}{\hbar} \int_{-\infty}^{\infty} (\sqrt{E - V} - \sqrt{E}) dx. \tag{6.84}$$

The delay time is obtained by plugging this result into formula (6.80), which gives

$$\Delta t = \sqrt{2m} \frac{\partial}{\partial E} \int_{-\infty}^{\infty} \left(\sqrt{E - V} - \sqrt{E}\right) dx = \int_{-\infty}^{\infty} \left(\sqrt{\frac{m}{2(E - V)}} - \sqrt{\frac{m}{2E}}\right) dx, \tag{6.85}$$

i.e.

$$\Delta t = \int_{-\infty}^{\infty} \left(\frac{1}{v(x)} - \frac{1}{v_0}\right) dx, \tag{6.86}$$

where $v(x)$ is the classical velocity at x. This is the same result that classical theory gives: the delay time is the total travel time $t_v = \int_{-\infty}^{\infty} dx/v(x)$ in the presence of the potential minus the travel time $t_0 = \int_{-\infty}^{\infty} dx/v_0$ that the free particle would need. In other words, in this approximation, the potential slows down the packet as a whole, as if it were a rigid structure. This is a direct consequence of the assumption that the real part of the reduction factor $S(\omega)$ is the same for the whole package.

Several different definitions for parameters such as delay time, transmission time, and so on are known in the specialized literature, the one used here to arrive at Equation (6.86) is one of them.

6.3 Electronic Properties of Solids

The mechanical, acoustic, thermal, or electrical properties of a body, be it a gas, a liquid or an ordered or disordered solid, are strongly dependent on the atomic or molecular content and structure. In particular, the electrical properties of solids are essentially determined by the dynamics of the electrons in their interior, leading to their classification as conductors, semiconductors, or insulators, depending on their ability to transfer charge and energy.

https://doi.org/10.1017/9781009679633.007 Published online by Cambridge University Press

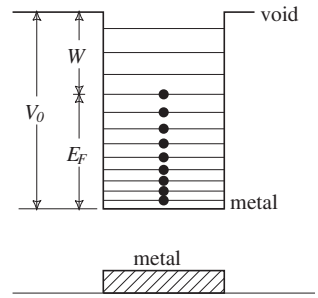


Figure 6.5 Elementary model of a metal. E_F is the Fermi energy, determined by the highest occupied level; W is the work function, and V_0 is the depth of the well. At $T = 0$ each of the lower levels is occupied by two electrons.

We begin the second part of this chapter by considering a simple model for conduction electrons in a metal, and continue with a discussion of the physics of electrons in a periodic structure, which allows us to explain important electronic properties of conductors and semiconductors and various combinations thereof, which form the basis of most of today's electronics.

6.3.1 The free-electron-gas model

About 80% of the 118 elements in the modern periodic table are metals.[14] The metals usually have 1, 2, or 3 electrons in their valence shell (i.e., the outer shell, see what follows), so they can donate electrons very easily. The high (electrical and thermal) conductivity of metals is due to the large number – of the order of 10^{22} /cm^3 – of electrons that can move (almost) freely in the material. This allows us to think of these conduction electrons as a "free-electron gas" inside the metal (a brief explanation of the physical origin of this phenomenon is given in Chapter 15). However, these electrons cannot escape from the material because work is required to extract them (the so-called work function W of the metal), which suggests that the metal should be represented in a first approximation as a finite potential well, which for simplicity will be assumed to be rectangular in the following.

Our first task is to estimate the depth of the well, for which we proceed as follows. Due to the large number of atoms in the metal, there is a very large number of stationary energy levels within the well – as will be shown below – so the electronic energy scheme of the metal is similar to that shown in Fig. 6.5. In the ground state (i.e., at absolute zero temperature), the electrons trapped in the well fill the lowest energy levels; each level can be occupied by a maximum of two electrons, due to the Pauli exclusion principle, as we will have occasion to study in Section 15.2. This means that if the metal contains N free electrons, at $T = 0$ they will occupy the $N/2$ lowest energy levels, producing a rectangular distribution. As the temperature increases, the higher levels begin to fill up; the distribution is no longer rectangular and becomes smoother, as shown in Fig. 6.6. This is the *Fermi–Dirac distribution*, which we will discuss later.

[14] This fraction is not a well-defined quantity, as the boundary between metals and nonmetals varies due to the lack of universally accepted definitions.

https://doi.org/10.1017/9781009679633.007 Published online by Cambridge University Press

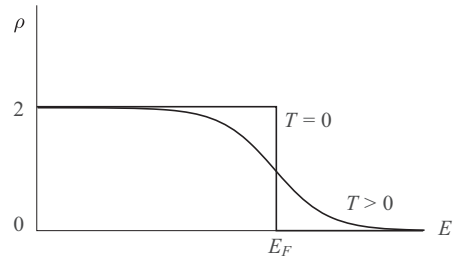


Figure 6.6 Electron level occupancy in a metal as a function of energy and temperature, according to the Fermi–Dirac distribution.

The highest energy level occupied determines the *Fermi energy* E_F,[15] and the surface in momentum space with $E = E_F$ is called the *Fermi surface*. To calculate the Fermi energy at zero temperature, we consider the metal as a cube with side L, so that the free electrons have momentum p_i $(i = x, y, z)$ given by

$$p_i = \hbar k_i = \frac{2\pi\hbar}{L} n_i, \quad n_i = 0, 1, 2, \ldots n_{i\,\max}. \tag{6.87}$$

Increasing n_i by one unit will give the smallest possible increase in momentum. Since, for a piece of metal, L is very large compared to the size of the atoms, this change in momentum is very small, so we can write $\Delta n_i \to L dp_i/2\pi\hbar$. The number of states contained in the interval $(p, p + dp)$ is therefore

$$\Delta^3 n = \Delta n_x \Delta n_y \Delta n_z \to \frac{L^3}{(2\pi\hbar)^3} d^3 p. \tag{6.88}$$

Since $2\Delta^3 n$ electrons fit into each interval, if $p_{\max}$ is the momentum of the electrons on the Fermi surface, the total number N of free electrons becomes

$$\begin{aligned} N &= 2\sum \Delta n_x \Delta n_y \Delta n_z \to \frac{2L^3}{(2\pi\hbar)^3} \int_0^{p_{\max}} d^3 p \\ &= \frac{L^3}{4\pi^3\hbar^3} 4\pi \int_0^{p_{\max}} p^2 dp = \frac{L^3 p_{\max}^3}{3\pi^2\hbar^3}, \end{aligned} \tag{6.89}$$

and for the density ρ_0 of free electrons in the metal we get $\rho_0 \equiv N/V = p_{\max}^3/3\pi^2\hbar^3$, hence

$$p_{\max} = \hbar(3\pi^2\rho_0)^{1/3}. \tag{6.90}$$

It follows that the Fermi energy is

$$E_F = \frac{p_{\max}^2}{2m} = \frac{\hbar^2}{2m}(3\pi^2\rho_0)^{2/3}. \tag{6.91}$$

[15] Enrico Fermi (1901–1954) was an exceptional Italian–US theoretical and experimental physicist, one of the great physicists of the twentieth century. He made a long list of contributions to quantum theory, nuclear physics, elementary particles, and statistical physics, and developed the first nuclear reactor (pile) in Chicago. He received the Nobel Prize in 1938 for his discovery of induced radioactivity. The (artificial) radioactive element Z=100 is named after him as *fermium*.

https://doi.org/10.1017/9781009679633.007 Published online by Cambridge University Press

This allows us to estimate the Fermi energy for a representative metal. For example, assuming that there is one free electron per atom in silver (which is a realistic assumption), we can calculate the electron density as the number of silver atoms per unit volume. Since silver has a density of 10.5 g cm^{-3} and an atomic weight of 107.9 u (unified atomic mass units), with Avogadro's number (number of atoms in one gram-atom) equal to 6.022×10^{23}, we get

$$\rho_0 \approx \frac{10.5}{107.9} \times 6.022 \times 10^{23} \approx 6 \times 10^{22} \text{ cm}^{-3}. \tag{6.92}$$

Substituting this value into Eq. (6.91), we get $E_F \approx 8 \times 10^{12}$ erg ≈ 5.5 eV. Since the work function of silver is 4.7 eV, the depth of the potential well is 10.2 eV (see Fig. 6.5).

Equation (6.91) allows us to understand another property of metals that is not classically explained. When the electron-gas theory emerged as a model for metals, it was considered, based on classical arguments, that each electron contributes an amount $\frac{3}{2}k$ (k is Boltzmann's constant) to the specific heat, due to the equipartition theorem. However, experimentally it is observed that the free electrons do not contribute to the specific heat of the metal. This contradiction is easily resolved by the quantum theory just formulated. In fact, we can roughly calculate the average energy of the free electrons at low temperatures as

$$\overline{E}_{\text{elect}} = \frac{2}{N} \sum E(n) \Delta^3 n \to \frac{2}{\rho_0 L^3} \int \frac{p^2}{2m} \frac{d^3 p}{(2\pi\hbar)^3} L^3 = \tfrac{3}{5} E_F. \tag{6.93}$$

As the temperature rises, the electrons with an energy close to (but less than) E_F start to absorb energy and populate the upper levels close to E_F one by one. At low temperatures, the energy corrections due to these absorptions are negligible, which means that their contribution to the specific heat of the metal is effectively zero,

$$c_v^{\text{elect}} \equiv \left(\frac{\partial \overline{E}_{\text{elect}}}{\partial T} \right)_{T=0} = 0. \tag{6.94}$$

A calculation at higher temperatures allows us to verify that $c_v^{\text{elect}} \ll \frac{3}{2}k$, because the electron-level occupancy is determined by the Fermi–Dirac distribution, as is illustrated in Fig. 6.6 and will be discussed in Chapter 15.

6.3.1.1 Electron tunneling in metals

The electron-gas model can also explain various processes of electron emission from a metal. In general, it is possible to extract electrons by adding some extra energy to the metal. If this is done by increasing the temperature of the metal, thermionic emission occurs. This method is used to release electrons from an electrode (usually a filament) in various electronic devices, such as vacuum tubes, and also for power generation or cooling. For significant thermionic emission to occur, the metal must be heated to a temperature above 10^3 K to overcome its work function W.

It is also possible to extract electrons by applying a sufficiently strong uniform external electric field (*cold* or *field emission*). In this case, the potential energy diagram is similar to that shown in Fig. 6.7, which consists of the superposition of the metal's potential well of depth V_0 and the external potential $-e_0 \mathcal{E} x$, where $\mathcal{E}$ is the electric

https://doi.org/10.1017/9781009679633.007 Published online by Cambridge University Press

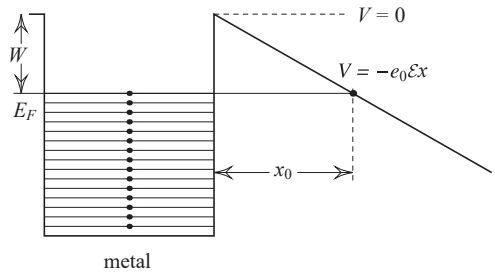


Figure 6.7 Field emission: the tunneling of electrons from a metal surface in the presence of an electric field. Tunneling occurs mainly with electrons that have an energy $E \simeq E_F$.

field strength. This creates a triangular potential barrier through which electrons can escape. The width of the barrier, x_0, is given by the intersection at $W = e_0\mathcal{E}x_0$, because at normal temperatures the tunneling electrons have an energy close to the Fermi level. Equation (6.56) thus gives for the transmission coefficient

$$T = e^{-\frac{2}{\hbar}\sqrt{2m}\int_0^{x_0}\sqrt{W-e_0\mathcal{E}x}dx}. \tag{6.95}$$

that is,

$$T = e^{-\mathcal{E}_0/\mathcal{E}}, \tag{6.96}$$

where $\mathcal{E}_0$ is given by

$$\mathcal{E}_0 = \frac{4\sqrt{2m}W^{3/2}}{3\hbar e_0}. \tag{6.97}$$

An estimate of the value of this parameter under normal conditions is 10^6 V/cm. Equation (6.96) shows that fields of this intensity produce relatively strong emission currents.

Both thermionic and field emission had already been produced in the late nineteenth century, but could only be explained with the advent of QM. In fact, the quantum tunneling of electrons is considered one of the theory's early triumphs. Since then, electron tunneling has found a number of important applications. One example is the *scanning tunneling microscope* (STM), which is used to image surfaces at the atomic level. A very sharp conductive tip is brought close to the surface to be examined and a voltage is applied between the two, causing electrons to pass through the gap between them (see Fig. 6.8). The resulting tunnel current, which depends on the position of the tip and the applied voltage according to Eqs. (6.96) and (6.97), is monitored as the tip traverses the surface, and is usually displayed as an image.

6.3.1.2 The contact potential

Another example of the use of the simple metal model is the *contact potential*. This phenomenon, discovered by Volta,[16] occurs when the surfaces of two different metals

[16] Alessandro Volta (1745–1827) was an Italian physicist and chemist, inventor of the electrochemical battery (the common pile, 1799), and discoverer of methane gas. The unit of voltage of electric strength, the *volt*, is named after him.

https://doi.org/10.1017/9781009679633.007 Published online by Cambridge University Press

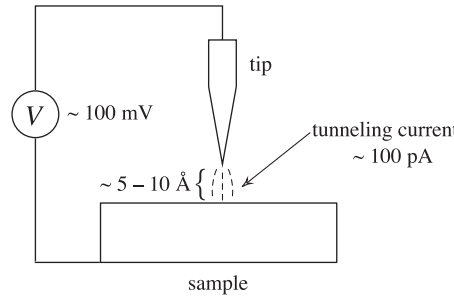


Figure 6.8 Schematic of a scanning tunneling microscope.

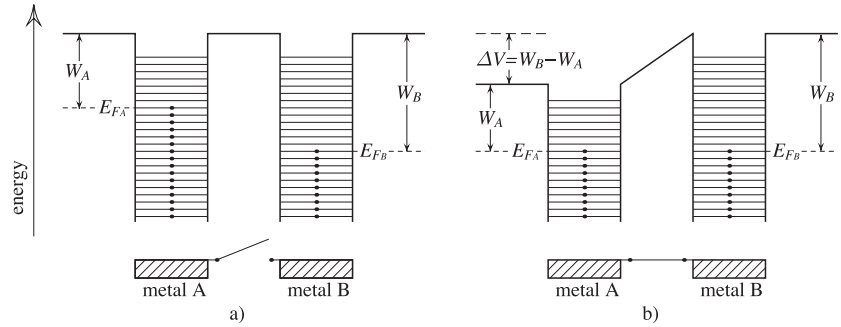


Figure 6.9 (a) Metals A and B having different work functions W_A, W_B. (b) The metals are brought into electrical contact Tunneling occurs from A to B until the Fermi levels are equalized.

come into contact, creating an electric potential difference between them. As long as the two metals are electrically isolated, the wells representing them will be at the same upper level, corresponding to the (common) vacuum potential, as shown in Fig. 6.9(a). When brought into electrical contact, the free electrons of each metal are separated from those of the other by a potential barrier. If the Fermi energy of metal A is higher than that of metal B, the more energetic electrons of metal A will tunnel into B, because there are holes in B for them to occupy. Since the reverse process does not occur, the net effect is a transfer of electrons from metal A to metal B, until the Fermi levels of the two metals are equalized, as shown in Fig. 6.9(b).

The number of electrons transferred is a very small fraction of the total number of free electrons, so that the Fermi energies are not significantly altered; therefore, the potential difference created between the metals is essentially given by the difference in work functions,

$$\Delta V = W_B - W_A. \tag{6.98}$$

Note that the explanation given is independent of the assumed shape of the well; in other words, the result (6.98) is essentially independent of the specific model.

In a chemical cell (an elementary battery) made of different metals, the potential difference between the electrodes is essentially equal to the corresponding contact potential given by (6.98), which also determines the direction of the current; the energy is extracted from the chemical processes taking place at the electrodes, involving the ions in the electrolyte.

https://doi.org/10.1017/9781009679633.007 Published online by Cambridge University Press

The work function of a metal changes slightly with temperature, so the contact potential also changes with temperature. Thus, when two strips of different metals are joined at their ends to form a ring (a device that constitutes a *thermocouple*) and one of the junctions is heated, an electromotive force is generated, known as the thermoelectric potential; this phenomenon was discovered by J. Seebeck (1770–1831) in 1821. The inverse phenomenon discovered by J. Peltier (1785–1845) in 1834, which is also reversible, occurs when an electromotive force is applied to one of the junctions of the thermocouple: one of them cools down and the other heats up.

6.3.2 Conductors and semiconductors

Although the model of the metal as a simple potential well has allowed us to explain a number of important properties, it has the serious disadvantage that it takes no account of the internal structure of the material, which produces fundamental effects that cannot be understood without it. The physics of electrons in a solid is, strictly speaking, a many-body problem because the electrons interact with each other, yet most of the relevant features can be derived by assuming that they move independently in an effective potential. Therefore, to introduce us to the study of these effects, we will model the material as a periodic sequence of similar wells separated by barriers, with the center of each well corresponding to the position occupied by one of the positive ions in the lattice.

6.3.2.1 The Kronig–Penney model

To keep the study as simple as possible, we consider only the 1D case, with an infinite number of rectangular wells and barriers, as shown in Fig. 6.10. This periodic single-electron model, proposed and studied by R. Kronig (1904–1995) and W. G. Penney (1909–1991) in 1930, has an analytical solution, as we shall see, and therefore allows simple calculations.

The Kronig–Penney potential is defined by three parameters: the lattice periodicity a, the well's width b and its height V_0. The position of a point x in well n is given by the coordinate x_n measured from the left internal wall of this well,

$$x_n = x - na = x \bmod a, \tag{6.99}$$

and similarly for the neighboring wells. Note that point B has the coordinate $x_n = a$, which is equivalent to $x_{n+1} = 0$. The wave function inside the well n can be written as

$$\psi_n = A_n \sin kx_n + B_n \cos kx_n, \quad 0 \le x_n \le b, \tag{6.100}$$

while inside the barrier n it is

$$\psi_n = A'_n \sinh q(a - x_n) + B'_n \cosh q(a - x_n), \quad b \le x_n < a, \tag{6.101}$$

with k and q given by Eqs. (5.2) and (5.3), respectively. Likewise, the wave function inside the well $n + 1$ is evidently

$$\psi_{n+1} = A_{n+1} \sin kx_{n+1} + B_{n+1} \cos kx_{n+1}, \quad 0 \le x_{n+1} \le b. \tag{6.102}$$

https://doi.org/10.1017/9781009679633.007 Published online by Cambridge University Press

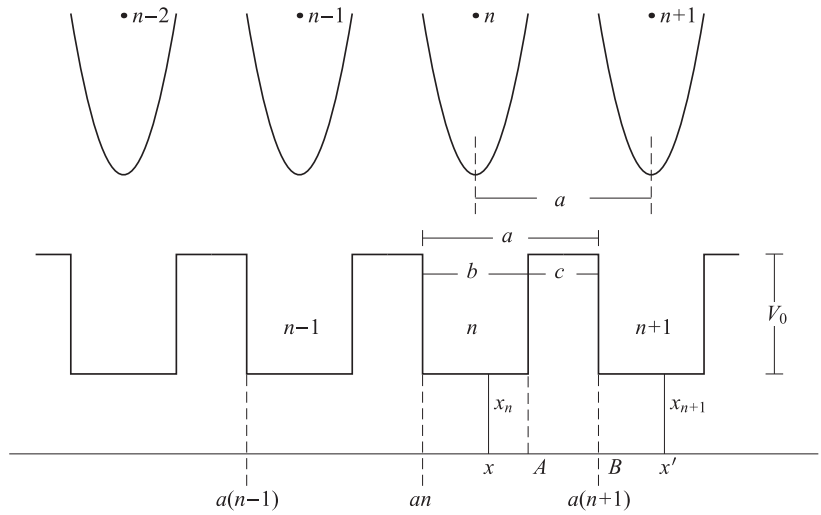


Figure 6.10 The Kronig–Penney model. Electrons move in a one-dimensional crystal under the influence of a periodic potential $V(x) = V(x + a)$, where a is the lattice parameter. In the lower diagram, b is the width of the well and $c = a - b$ is the width of the barrier of height V_0.

The continuity conditions at point A give ($x_n = b$; $a - b = c$)

$$A_n \sin kb + B_n \cos kb = A'_n \sinh qc + B'_n \cosh qc, \tag{6.103}$$

$$k[A_n \cos kb - B_n \sin kb] = -q[A'_n \cosh qc + B'_n \sinh qc]. \tag{6.104}$$

Analogously, the continuity conditions at point B give ($x_n = a$, $x_{n+1} = 0$)

$$B'_n = B_{n+1}, \quad -qA'_n = kA_{n+1}. \tag{6.105}$$

We use these two expressions to eliminate A'_n and B'_n from Eqs. (6.103) and (6.104), thus obtaining

$$A_n \sin kb + B_n \cos kb = -(k/q)A_{n+1} \sinh qc + B_{n+1} \cosh qc, \tag{6.106}$$

$$A_n \cos kb - B_n \sin kb = A_{n+1} \cosh qc - (q/k)B_{n+1} \sinh qc. \tag{6.107}$$

To simplify the solution of this system of equations, we take the limit in which $V_0 \to \infty$ and at the same time the barriers become narrower and narrower, $c \to 0$, with the product cV_0 remaining constant (this limit is essentially that of a periodic succession of repulsive delta potentials). We write this constant in the form

$$cV_0 = \frac{\hbar^2 \lambda}{ma}, \tag{6.108}$$

where λ is a (real and positive) dimensionless parameter. In this limit, $kb = ka$, $qc = \sqrt{2m\lambda c/\hbar^2} \to 0$ and $q^2 c = 2\lambda/a$. On the other hand, since qc is very small, we can write $\sinh qc \simeq qc$ and $\cosh qc \simeq 1$. With all these approximations, Eqs. (6.106) and (6.107) reduce to

$$A_n \sin ka + B_n \cos ka = B_{n+1}, \tag{6.109}$$

https://doi.org/10.1017/9781009679633.007 Published online by Cambridge University Press

$$A_n \cos ka - B_n \sin ka = A_{n+1} - \frac{2\lambda}{ka} B_{n+1}. \tag{6.110}$$

An inspection of this system of equations leads one to propose a general solution of the form

$$A_n = C_1 f^n, \quad B_n = C_2 f^n, \tag{6.111}$$

where f is a function to be determined and C_1 and C_2 are two constants, also to be determined. Substituting these expressions in the previous system, one obtains

$$C_1 \sin ka + C_2(\cos ka - f) = 0, \tag{6.112}$$

$$C_1(\cos ka - f) - C_2 \left(\sin ka - \frac{2\lambda}{ka} f \right) = 0. \tag{6.113}$$

For solutions with $C_1, C_2 \neq 0$ to exist, the determinant of the system must vanish. As a result, we get

$$f^2 - 2\left(\cos ka + \frac{\lambda}{ka} \sin ka \right) f + 1 = 0. \tag{6.114}$$

At this point it is convenient to define a new (real or complex) parameter $\varkappa$ such that

$$\cos \varkappa a \equiv \cos ka + \frac{\lambda}{ka} \sin ka, \tag{6.115}$$

in terms of which Eq. (6.114) takes the form

$$f^2 - 2f \cos \varkappa a + 1 = 0. \tag{6.116}$$

The solutions of this equation, $f = e^{\pm i \varkappa a}$, substituted into Eqs. (6.112) and (6.113), give for C_1 and C_2

$$C_1 = A \frac{f - \cos ka}{\sin ka}, \quad C_2 = A, \tag{6.117}$$

where A is the normalization constant. The double sign of the solutions indicates that the electrons can drift to the right and to the left. Taking (without loss of generality) only the positive sign, we determine A_n and B_n with the help of Eqs. (6.111); this gives

$$A_n = A \frac{f - \cos ka}{\sin ka} e^{i \varkappa na}, \quad B_n = A e^{i \varkappa na}. \tag{6.118}$$

Substituting these values into Eq. (6.100) for ψ and writing the product na in terms of the difference $x - x_n$, Eq. (6.120), we obtain, after some elementary trigonometric transformations,

$$\begin{aligned} \psi_n &= A \frac{f - \cos ka}{\sin ka} e^{i \varkappa na} \sin kx_n + A e^{i \varkappa na} \cos kx_n \\ &= A e^{i \varkappa x} \frac{e^{i \varkappa (a - x_n)} \sin kx_n + e^{-i \varkappa x_n} \sin k(a - x_n)}{\sin ka}, \end{aligned} \tag{6.119}$$

that is,

$$\psi_n = A e^{i \varkappa x} u_n, \tag{6.120}$$

where u_n is given by

$$u_n = \frac{e^{i \varkappa (a - x_n)} \sin kx_n + e^{-i \varkappa x_n} \sin k(a - x_n)}{\sin ka}. \tag{6.121}$$

https://doi.org/10.1017/9781009679633.007 Published online by Cambridge University Press

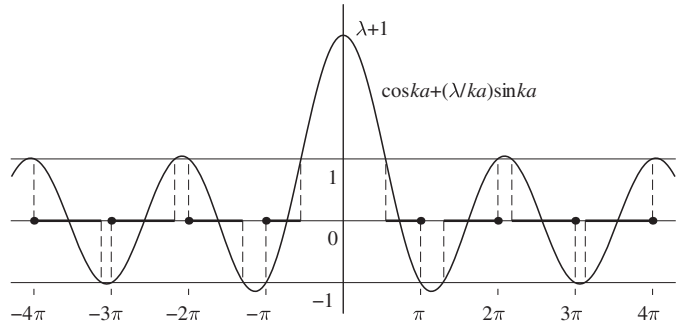


Figure 6.11 Graphical determination of the solutions of Eq. (7.92). Electrons with κ real, i.e., $|\cos \kappa a| \leq 1$, can move freely through the crystal; the regions corresponding to these allowed values of *ka* are marked with thick lines on the graph. As the barrier height increases, the allowed regions become narrower, until they disappear in the limit $\lambda \to \infty$ (thick dots on the graph, corresponding to $ka = n\pi$).

The fundamental equation (6.120), known as the *Bloch wave function*, was obtained here for the particular case of rectangular wells, but any periodic potential produces wave functions of this general form, a result that is known as *Floquet's theorem*. This theorem states that linear differential equations with periodic coefficients have solutions of the form $e^{\pm ikx}u(x)$, where $u(x)$ is a periodic function of x.[17] Bloch waves are the specific case applied to Schrödinger's equation, also when the periodic potential is not that of Kronig and Penney.

6.3.2.2 Electronic band structure

According to Eq. (6.120), when $\varkappa$ is real, the electrons move freely in the crystal. In this case the eigenfunction ψ_n given by Eq. (6.120) represents a plane wave times a periodic function, with the energy of the electrons determined by the value of $\varkappa$. To find the possible eigenvalues of the energy we have to study the solutions of Eq. (6.115) for real κ. This requires that ka is such that

$$-1 \leq \cos ka + \frac{\lambda}{ka} \sin ka \leq 1. \tag{6.122}$$

Figure 6.11 shows the function $\cos \varkappa a = \cos ka + (\lambda/ka) \sin ka$ for a given value of the parameter λ; the allowed values of ka (Brillouin zones) are marked with thick lines on the graph. We see that the solution defines a sequence of continuous bands, whose origin and position we can understand as follows. If for a moment we imagine the crystal as a set of wells isolated by impenetrable barriers (λ infinite), the solutions of the eigenvalue equation (6.115) would be precisely those corresponding to a single infinite well with a discrete energy spectrum, $ka = n\pi$, as can be seen in Fig. 6.12(a). By lowering the barriers or making them narrower (λ finite), the wells become interconnected and each energy level breaks down into N close levels, as we saw in Section 5.5. Since N is infinite in the model used, this decomposition produces an infinite number of infinitely close levels. This results in a series of bands, each with a continuum of

[17] A proof of Floquet's theorem can be found in Margenau and Murphy (2009), section 2.17.

https://doi.org/10.1017/9781009679633.007 Published online by Cambridge University Press

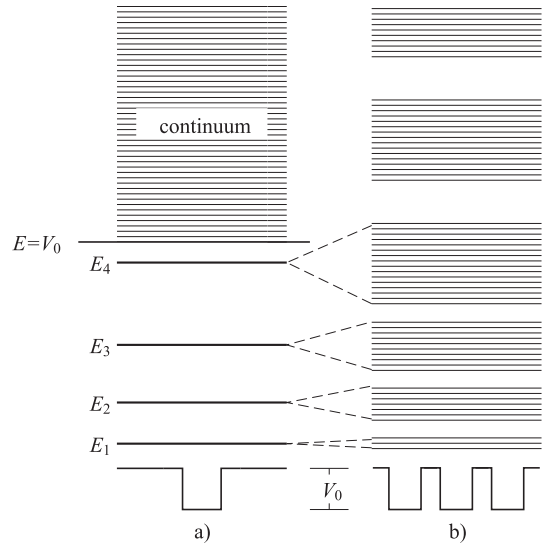


Figure 6.12 (a) Discrete energy levels for electrons in a single finite potential well for $E < V_0$ and continuous energy spectrum for $E > V_0$. (b) When the potential wells connect, continuous energy bands appear, separated by forbidden energy regions.

levels around the solution for an isolated well, as shown in Fig. 6.12(b). Note that the bands can also occur in regions that originally correspond to the continuum of one of the isolated wells.

In a real crystal, the bands are not strictly continuous, but are formed by a succession of discrete levels in close proximity, since the number of ions in the lattice is finite. The latter is the model we will use in the following to discuss qualitatively some important issues in conduction theory; more complete and quantitative analyses can be found in the literature on solid-state theory.

Consider a metal that contains one valence electron per atom (e.g., an alkali; the meaning of this term is discussed in Chapter 16). Since each electronic state can be occupied by up to two electrons, in the ground state the electrons occupy half the levels of the lower band (the band has N levels, where N is the number of atoms). Half of the electrons move in one direction and half in the opposite direction, so there is no net current. If we apply a voltage to the metal, more electrons will move in the opposite direction to the field; since the lower $N/2$ levels are already occupied, the electrons from the extra flow have to move to higher energy levels. Therefore, for electrical conduction to occur, the lower band must fulfil one of two conditions: a) have a sufficient number of vacancies (unoccupied levels), as in the alkali metal example, or b) be in contact with the second band. For this reason, the second band is called the *conduction band*, while the first is called the *valence band*.

In an insulator, on the other hand, all levels of the valence band are occupied, those of the conduction band are empty, and there is a separation of several eV between the two bands. This band gap is large enough to prevent electrons from passing into the conduction band when an external electric field is applied, so the material does not conduct.

https://doi.org/10.1017/9781009679633.007 Published online by Cambridge University Press

Note the essential role played in the preceding argument by the fundamental fact that each electron level can be occupied by a maximum of two electrons (with different spin states); without this limitation, all crystals would be conductors. We can therefore say that the basic theory of conduction provides an excellent verification of this fundamental principle (or, more generally, of the fact that electrons satisfy the Fermi–Dirac statistics, as we shall see in Chapter 15).

In some materials it happens, however, that the energy gap (the zone between the valence and conduction bands) is comparatively very narrow. For example, for Si it is 1.14 eV and for Ge only 0.67 eV. These crystals remain insulators at $T = 0$, but their conductivity increases rapidly with temperature, as some thermally excited electrons pass into the conduction band. At room temperature ($kT \approx 10^{-2}$ eV) these materials reach conductivities intermediate between those of a dielectric and a conductor, and are therefore generically called *semiconductors*. Electrons that pass into the conduction band leave vacancies in the valence band, called *holes*.

The intrinsic conductivity of a semiconductor, that due to thermal excitation, has two contributions. On the one hand, electrons in the conduction band can produce a current in the usual way; on the other hand, when an electron in the valence band occupies one hole, but leaves another, a migration of vacancies occurs. It can be shown that this movement of electrons in a nearly full valence band (with few holes) can be consistently considered as the migratory movement – in the opposite direction – of the holes themselves. The current due to the movement of electrons is called *n-type*, while that due to the movement of vacancies is called *p-type*.

6.3.2.3 Extrinsic semiconductors

The conductivity of a semiconductor can be increased in various ways. For example, photons in the red or infrared range already have sufficient energy to transfer electrons into the conduction band; that is, some semiconductors are *photoconductors*. This property is widely used in the manufacture of photosensitive semiconductor devices.

A more important mechanism for increasing the conductivity of a semiconductor is to add impurities (or *dopands*) to the crystal lattice. For example, suppose that a small amount of arsenic is added during the preparation of a germanium crystal. Since As has five valence electrons compared to Ge's four, when an As atom takes the place of a Ge atom in the lattice, one valence electron is (almost) free and is easily transferred to the conduction band by thermal excitation. A significant effect on the conductivity of the material due to these substitutions is achieved at densities as low as one impurity per 10^6 atoms. Impurities that contribute electrons to the conduction band are called *donors* and the resulting semiconductor is *n*-type (with an excess of free electrons in the conduction band).

Impurities can also be introduced to create an electron deficit (presence of vacancies) in the valence band. These impurities are called *acceptors* and the resulting semiconductor is called *p*-type (with an excess of holes in the valence band). An example of this is when Ga is added to Ge; as Ga has only three valence electrons, it creates a hole in the band.

The energy levels produced by donors and acceptors are in the forbidden zone, very close to the conduction or valence band, respectively, as illustrated in Fig. 6.13.

https://doi.org/10.1017/9781009679633.007 Published online by Cambridge University Press

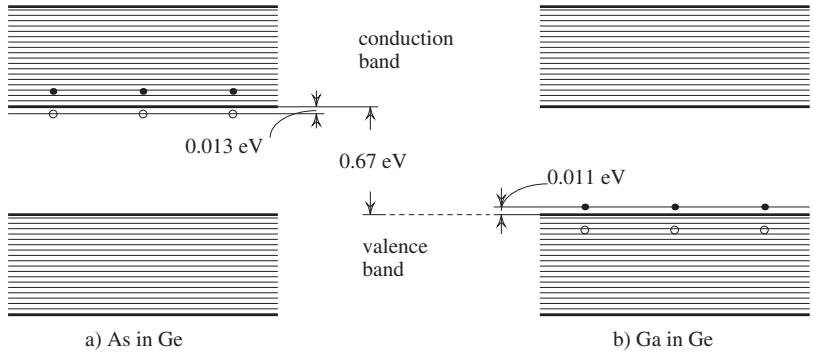


Figure 6.13 The valence and conduction bands of a doped Ge crystal. (a) Energy levels of a donor impurity (As) appear close to the conduction band (n-type doping); (b) energy levels of an acceptor impurity (Ga) appear close to the valence band (p-type doping).

The properties of semiconductors can vary over a wide range by controlling the nature and proportion of added impurities. It is these fundamental principles that are used in the construction of a variety of semiconductor electronic devices such as diodes, transistors, rectifiers, modulators, photocells, integrated circuit chips, in addition to LEDs and semiconductor lasers.

A semiconductor *diode* is formed by joining a p-type semiconductor with an n-type semiconductor; this is called a p–n junction. At the junction, electrons accumulate in the p region and holes accumulate in the n region, which is why it has some similarities to a charged capacitor. The device has resistances of very different orders of magnitude for currents flowing in one direction or the other, so it behaves like a rectifier (an element that allows electric current to flow in only one direction).

A *tunnel diode* is a p–n junction with a high concentration of impurities; in this case, the situation is such that applied voltages can modify and control the passage of electrons from one side of the junction to the other, so that for certain external voltages the net current decreases with the voltage, i.e., the device has negative resistance. This feature makes the tunnel diode extremely useful in the construction of many different types of electronic generators and switching circuits, mainly because of its very fast response. For example, it can be used to make generators that operate at frequencies above 10^{11} Hz.

The LED (acronym for *light-emitting diode*) is an electroluminescent heavily doped p-n junction diode. When current flows through the diode, electrons recombine with holes at the junction, releasing energy in the form of light, the color of which is determined by the energy band gap of the semiconductor used.

A *transistor* is obtained by juxtaposing two junctions, creating an n–p–n or p–n–p combination. These devices are used to control currents, so they can act as power, current or voltage amplifiers, as oscillators, and so on. Their great versatility, economy, smallness, durability, low working voltages, and so on, have made them indispensable in contemporary electronic technology, displacing thermionic tubes (bulbs) and allowing the construction of ultra-compact circuits (as, for example, in modern computers and cell phones).

https://doi.org/10.1017/9781009679633.007 Published online by Cambridge University Press

Recent attention to low-dimensional systems that have electrons and holes confined to move in a plane, along a wire or within a quantum dot, forming structures with submicron resolution, has given rise to new physical concepts and applications, distinct from their three-dimensional analogs. Interested students are recommended to consult the specialized bibliography for more information on the subject.

Problems

P6.1 Derive Eq. (6.56) for the transmission coefficient T from the expression of the wave function for a rectangular barrier obtained in Chapter 5, considering the barrier of arbitrary profile as built up by a succession of rectangular barriers that become infinitely narrow in the limit.

P6.2 A particle moves in a conical potential that vanishes for $|x| > a$, has the value $-V_0$ at $x = 0$ and goes linearly from $-V_0$ to 0 in the intervals $|x| \leq a$. Study the bound states of the system using the WKB method.

P6.3 An electron with energy $E > V_0$ moves in the potential

$$V(x) = V_0 \cosh^{-2}(x/a), \quad V_0 > 0. \tag{6.123}$$

Determine the conditions under which the WKB method is applicable and obtain the transmission coefficient through the barrier.

P6.3 Use the WKB method to find the quantization conditions for particles moving in a well, one of the walls of which is impenetrable.

P6.4 Use the WKB method to find the energy of the stationary states of particles moving inside a box with impenetrable walls. Analyze the results and compare them with the exact ones.

P6.5 Determine the transmission coefficient for a monochromatic electron beam of energy 1000 eV and intensity 1 mA incident on a rectangular barrier of width 10 Å and height of 200 eV. The barrier height is now increased to 2000 eV; determine the new transmission coefficient and compare the results.

P6.6 Use the WKB method to determine the energy levels of a one-dimensional harmonic oscillator. Discuss your result.

P6.7 Use the WKB method to find the ratio between the width and depth of a one-dimensional square well that is required for no bound state to exist.

P6.8 Use the semiclassical approximation to determine the average value of the kinetic energy of a stationary state.

P6.9 Apply the reasoning used to arrive at Eq. (6.91), to obtain the formulae for the density of states as a function of the maximum energy level:

a) for a 2D layer,
b) for a 1D wire.

P6.10 Consider an infinite 1D periodic potential $V(x) = V(x+a)$. Starting from the Schrödinger equation for ψ_k as an eigenfunction corresponding to the energy E_k, obtain the Schrödinger equation for $u_k(x)$.

https://doi.org/10.1017/9781009679633.007 Published online by Cambridge University Press

P6.11 Find the energy bands for the periodic potential

$$V(x) = g \sum_{n=-\infty}^{\infty} \delta(x - na), \quad g, a > 0. \tag{6.124}$$

P6.12 Estimate the minimum energy of the first allowed band for the previous problem, when $\lambda = 8$ and $g/a = 1$ eV.

Bibliographical Notes

For a detailed derivation of the WKB connection formulas, see Langer (1937), p. 669.
For a proof of Floquet's theorem, see Margenau and Murphy (2009), section 2.17.
Introductory studies on low-dimensional quantum systems can be found in Harrison (2000).
For a detailed treatment of semiconductors, see for example, Grundmann (2021).

https://doi.org/10.1017/9781009679633.007 Published online by Cambridge University Press