

CHAPTER 1

A BRIEF REVIEW OF THE RELEVANT BASIC EQUATIONS OF PHYSICS

From a mathematical viewpoint, all equations of physics (both microscopic and macroscopic) are relevant for semiconductor devices. In an absolutely strict mathematical way, we therefore would have to proceed from the fundamentals of quantum field theory and write down the $\approx 10^{23}$ coupled equations for all the atoms in the semiconductor device. Then we would have to solve these equations, including the complicated geometrical boundary conditions. However, the outcome of such an attempt is clear to everyone who has tried to solve only one of the 10^{23} equations.

Any realistic approach oriented toward engineering applications has to proceed differently. Based on the experience and investigations of many excellent scientists in this field, we neglect effects that would only slightly influence the results. In this way many relativistic effects become irrelevant. In my experience, the spin of electrons plays a minor role in the theory of most current semiconductor devices and can be accounted for in a simple way (the correct inclusion of a factor of 2 in some equations).

Most effects of statistics can be understood classically, and we will need only a very limited amount of quantum statistical mechanics. This leaves us essentially with the Hamiltonian equations (classical mechanics), the Schrödinger equation (quantum effects), the Boltzmann equation (statistics), and the Maxwell equations (electromagnetics).

It is clear that the atoms that constitute a solid are coupled, and therefore the equations for the movement of atoms and electrons in a solid are coupled. This still presents a major problem, a many-body problem. We will see, however, that there are powerful methods to decouple the equations and therefore make single particle solutions possible. The many interacting electrons in a solid are then, for example, replaced by single independent electrons moving in a periodic potential. Complex many body effects, such as superconductivity, are then excluded from our treatment, which is justified because of the low electron density in typical semiconductors. We also exclude in our treatment effects of extremely high magnetic fields because these are unimportant for most device applications.

In this way, the fundamental laws of physics are finally reduced to laws of semiconductor devices that are tractable and whose limitations are clearly stated. The following sections are written with the intent to remind the reader of the basic physics underlying device operation and to review some of the physicist's tool kit in solid-state theory.

1.1 THE EQUATIONS OF CLASSICAL MECHANICS, APPLICATION TO LATTICE VIBRATIONS

Hamilton was able to give the laws of mechanics a very elegant and powerful form. He found that these laws can be closely linked to the sum of kinetic and potential energy written as a function of momentumlike (p_i) and spacelike (x_i) coordinates.

This function is now called the *Hamiltonian function* $H(p_i, x_i)$. The laws of mechanics are

$$\frac{dp_i}{dt} = -\frac{\partial H(p_i, x_i)}{\partial x_i} \quad (1.1)$$

and

$$\frac{dx_i}{dt} = \frac{\partial H(p_i, x_i)}{\partial p_i} \quad (1.2)$$

where t is time and $i = 1, 2, 3$. Instead of x_i , we sometimes denote the space coordinates by x, y, z .

Some simple special cases can be solved immediately. The free particle (potential energy = zero) moves according to

$$H = \sum_i p_i^2 / 2m$$

and we have from Eq. (1.1)

$$\frac{dp_i}{dt} = 0; \quad p_i = \text{constant},$$

which is Newton's first law of steady motion without forces.

If we have a potential energy $V(x_1)$ that varies in the x_1 direction, we obtain from Eq. (1.1)

$$\frac{dp_1}{dt} = -\frac{\partial V(x_1)}{\partial x_1} \equiv F_o \quad (1.3)$$

The quantity defined as F_o is the force, and Eq. (1.3) is Newton's second law of mechanics.

A more involved example of the power of Hamilton's equations is given by the derivation of the equations for the vibrations of the atoms (or ions) of the crystal lattice. As we will see, these vibrations are of utmost importance

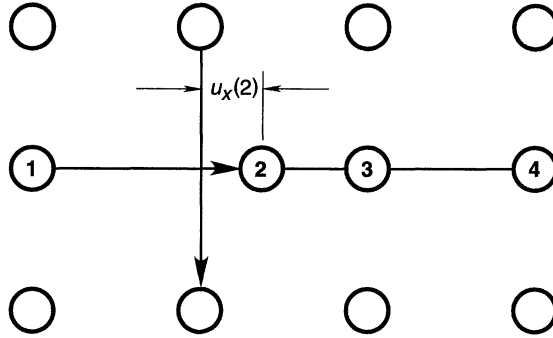


Figure 1.1 Displacement $u_i(r)$ of atoms in a crystal lattice.

in describing electrical resistance. They also give a fine example of how the many-body problem of atomic motion can be reduced to the solution of a single differential equation by using the crystal symmetry (group theory from a mathematics point of view) by a cut-off procedure. Here we cut off the interatomic forces beyond the nearest neighbor interaction. We also introduce below cyclic boundary conditions, which are of great importance and convenience in solid-state problems.

It suffices for this section to define a crystal, and we will be mostly interested in crystalline solids, as a regular array of atoms hooked together by atomic forces. “Regular” means that the distance between the atoms is the same throughout the structure. Many problems involving lattice vibrations can be solved by classical means (i.e., using the Hamiltonian equations) because the atoms that vibrate are very heavy. Then we only have to derive the kinetic and potential energy. Because we would like to describe vibrations (i.e., the displacement of the atoms), we express all quantities in terms of the atomic displacements $u_i(r)$, where $i = x, y, z$ and r is the number (identification) of the atom. It is important to note that r is not equal to the continuous space coordinate r in this chapter, although it has similar significance because it labels the atoms. The displacement of an atom in a set of regularly arranged atoms is shown in Figure 1.1.

We follow the derivations in Landsberg [5] and express the kinetic energy T by

$$T = \frac{1}{2} M \sum_n \dot{u}_i^2(r) \quad \text{where} \quad \dot{u}(r) = \frac{\partial u_i(r)}{\partial t} \quad (1.4)$$

and M is the mass of the atoms (ions).

We assume now that the total potential energy U of the atoms can be expressed in terms of a power series in the displacements,

$$U = U_0 + \sum_n B_i^r u_i(r) + \frac{1}{2} \sum_{\substack{rs \\ ij}} B_{ij}^{rs} u_i(r) u_j(s) + \cdots \quad (1.5)$$

where s also numbers the atoms as r does.

The following results rest on this series expansion and truncation, which makes a first principle derivation (involving many body effects) unnecessary.

Equation (1.5) is, of course, a Taylor expansion with

$$B_i^r = \frac{\partial U}{\partial u_i(r)} \quad (1.6)$$

Because the crystal is in equilibrium, that is, at a minimum of the potential energy U , the first derivative vanishes and

$$B_i^r = 0 \quad (1.7)$$

We further have

$$B_{ij}^{rs} = \frac{\partial^2 U}{\partial u_i(r) \partial u_j(s)} = B_{ji}^{sr} \quad (1.8)$$

We now use the fact that the crystal is translationally invariant—that is, we can shift the coordinate system by s atoms (start to count s atoms later), and the crystal is transformed into itself (at least if it is infinite). Therefore,

$$B_{ij}^{rs} = B_{ij}^{(r-s)0} \quad (1.9)$$

Furthermore, a rigid displacement (all u_i equal) of the crystal does not change U and we therefore have, from Eqs. (1.5) and (1.9),

$$\sum_r B_{ij}^{0r} = 0 \quad (1.10)$$

To derive Eq. (1.9), we have assumed an infinite crystal. We also could have introduced so-called periodic or cyclic boundary conditions; that is, continue the crystal by repeating it over and over. In one dimension, this means we consider only rings of atoms (Figure 1.2). This approach amounts to neglecting any surface effects or other effects that are sensitive to the finite extension of crystals.

We can now derive the equations of motion by using Eqs. (1.1) and (1.2) with coordinates $u_i(r)$ instead of x_i :

$$\dot{p}_i(r) = - \frac{\partial H(p_i, u_i)}{\partial u_i} \quad (1.11)$$

and

$$p_i(r) = M \dot{u}_i(r) \quad (1.12)$$

Eq. (1.11) gives

$$\dot{p}_i(r) = - \frac{1}{2} \frac{\partial}{\partial u_i(r)} \left[\sum_{\substack{mn \\ ij}} B_{ij}^{mn} u_i(m) u_j(n) \right] \quad (1.13)$$

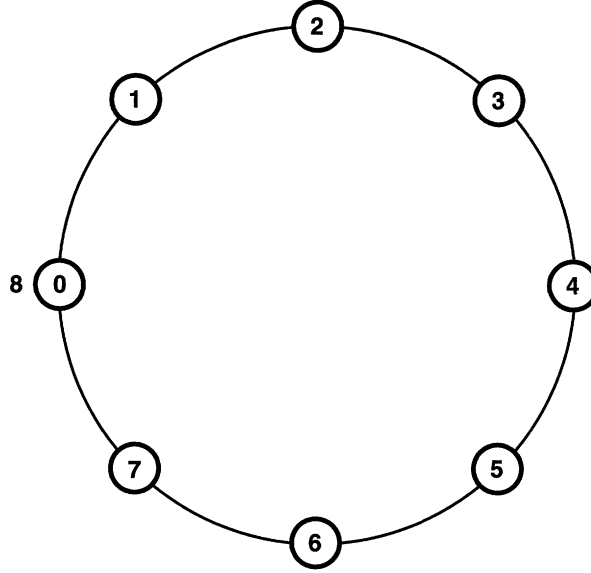


Figure 1.2 A ring of atoms representing cyclic boundary conditions.

Here, also, the indices m, n are used to number the atoms (as r, s above). Therefore

$$\dot{p}_i(r) = - \sum_{s,j} B_{ij}^{rs} u_j(s)$$

and, together with Eq. (1.12), one obtains

$$M \ddot{u}_i(r) + \sum_{s,j} B_{ij}^{rs} u_j(s) = 0 \quad (1.14)$$

Remember that the index s in Eq. (1.14) runs over a large number of atoms; that is, up to about 10^{23} in a typical crystal. The r can also assume any of these numbers. In other words, we have about 10^{23} coupled equations to solve. This situation is very typical for any type of solid-state problem, but by far not as hopeless as it may seem. Powerful methods have been developed to reduce the number of equations and the following treatment is representative. We will assume for simplicity that the crystal is one dimensional and avoid the complicated geometrical arrangement of atoms in a real crystal. (We will learn more about this when we discuss the electrons and their motion in crystals.)

In the three-dimensional case, Eqs. (1.7) through (1.10) are very helpful; they reduce the numbers of parameters. Without going into details, we mention that this reduction of parameters is generally accomplished by group theoretical arguments, and Eq. (1.9) is a direct consequence of the translational invariance (group of translations).

To proceed explicitly with our one-dimensional model, we need to make the drastic assumption that each atom interacts only with its nearest neighbor. (We can use the same method also for second, third . . . nearest neighbor interaction, if we proceed numerically and use a high-speed computer.) Our assumption means

$$B^{r,s} \neq 0 \quad \text{only for} \quad s = r \pm 1$$

Notice that we dropped the indices i, j because this is a one-dimensional problem. Without any loss of generality, we may assume $r = 0$. Then we have

$$B^{0s} \neq 0 \quad \text{for} \quad s = \pm 1$$

and

$$B^{0s} = 0 \quad \text{otherwise} \quad (1.15)$$

Furthermore,

$$B^{01} = B^{-10}$$

according to Eq. (1.9) and

$$B^{-10} = B^{0(-1)}$$

according to Eq. (1.8). Therefore,

$$B^{01} = B^{0(-1)}$$

From Eq. (1.10) we obtain

$$B^{00} = -2B^{01} \quad (1.16)$$

It is now customary to denote $B^{01} = B^{0(-1)}$ by $-\alpha$ (α is the constant of the “spring” forces that hold the crystal together) and therefore B^{00} by 2α . The equation of motion, Eq. (1.14), then becomes for any r

$$M\ddot{u}(r) = -2\alpha u(r) + \alpha u(r-1) + \alpha u(r+1) \quad (1.17)$$

Eq. (1.17) leaves us still with 10^{23} coupled differential equations. However, these equations are now in *tridiagonal form*, all with coefficient α . Such a form can be reduced to one equation by skillful substitution. The substitution can be derived from Bloch’s theorem, which we will discuss later. It also can be guessed:

$$u(r) = ue^{iqra} \quad (1.18)$$

Note that the amplitude u is still a function of time. Here a is the distance between atoms (i.e., the lattice constant). Eq. (1.17) becomes

$$M\ddot{u}e^{iqra} = -2\alpha ue^{iqra} + \alpha ue^{iqra}e^{-iqa} + \alpha ue^{iqra}e^{iqa} \quad (1.19)$$

which gives

$$M\ddot{u} = \alpha u(2\cos qa - 2)$$

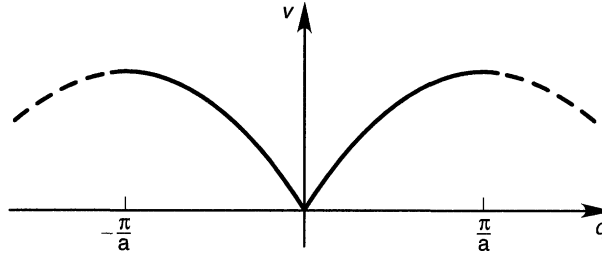


Figure 1.3 Dispersion relation $v(q)$ for lattice vibrations. (Remember, $E = \hbar v$.)

and

$$\ddot{u} = \frac{-4\alpha u}{M} \sin^2\left(\frac{aq}{2}\right)$$

This gives

$$\ddot{u} + v^2 u = 0 \quad (1.20)$$

with

$$v = 2 \left| \sin \frac{aq}{2} \right| \sqrt{\frac{\alpha}{M}} \quad (1.21)$$

This means that the atoms are oscillating in time with frequency v , which is a function of the wave vector q . The function is shown in Figure 1.3.

There are several important points to notice. First, at $q = -\pi/a$ and $q = \pi/a$, the energy has its highest value. For these q , the wavelength $\lambda = 2\pi/q$ has the value $\lambda = 2a$. As can be seen from Figure 1.4, this is the shortest wavelength that we really need to describe the physics of the lattice vibrations. Shorter wavelengths lead only to “wiggles” between the atoms, but the displacements are actually the same. For example, if $q = 3\pi/a$ and $\lambda = 2/3a$, the atoms are displaced in exactly the same way as for $q = \pi/a$. In other words, for any q outside the zone $-\pi/a \leq q \leq \pi/a$, which is called the *Brillouin zone*, we can find a q inside the zone that describes the same displacement, energy, and so

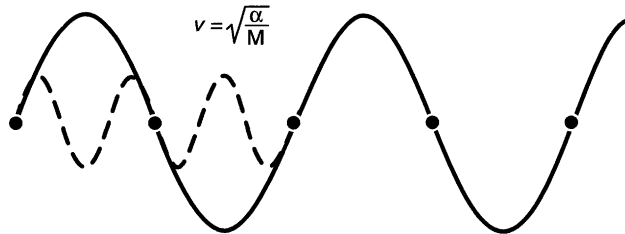


Figure 1.4 Illustration of the shortest possible physical wavelength of lattice vibrations.

on. Notice that in a real crystal the arrangement of atoms is different in different directions. Therefore, the three-dimensional Brillouin zone is usually a complicated geometrical figure (see Chapter 2).

Second, q is not a continuous variable because of the boundary conditions. Consider, for example, the ring of Figure 1.2 with eight atoms and

$$u(0) = u(8)$$

or, in general, for N atoms, we have

$$u(N) = u(0) \quad \text{and} \quad u(N) = u(0)e^{iqNa}$$

Therefore e^{iqNa} equals one, and we conclude that $q = 2\pi l/Na$ where l is an integer. If we restrict q to the first Brillouin zone, we have $-N/2 \leq l \leq N/2$. This means that q assumes only discrete (not continuous) values. However, because of the large number N , it can almost be regarded as continuous.

Third, without emphasizing it, we have developed a microscopic theory of sound propagation in solids. For small wave vectors q (i.e., large λ), we have

$$\sin \frac{qa}{2} \approx \frac{qa}{2}$$

and

$$v = \sqrt{\frac{\alpha}{M}} qa \quad (1.22)$$

Using $\lambda v = v_s$, where v_s is the velocity of sound, we obtain

$$v_s = 2\pi a \sqrt{\frac{\alpha}{M}} \quad (1.23)$$

which is a microscopic description of the sound velocity.

In real crystals additional complications arise from the fact that we can have two or even more different kinds of atoms. These atoms may oscillate as the identical atoms in the above example. There are, however, different modes of oscillation possible. If we think of a chain with two different kinds of atoms, it can happen that one kind of atom (black) oscillates against the other kind (white).

Such an oscillation can take place, and indeed does, at a very high (optical) frequency, and the corresponding lattice vibrations are called *optical phonons*. It is very important to note that in principle all black atoms can oscillate in phase against the white ones. This means that we can have high frequencies (energies) even if the wavelength is very large or the q vector is very small (Figure 1.5).



Figure 1.5 Two different kinds of atoms oscillating against each other. This represents a wave with high energy (frequency) and small wave vector.

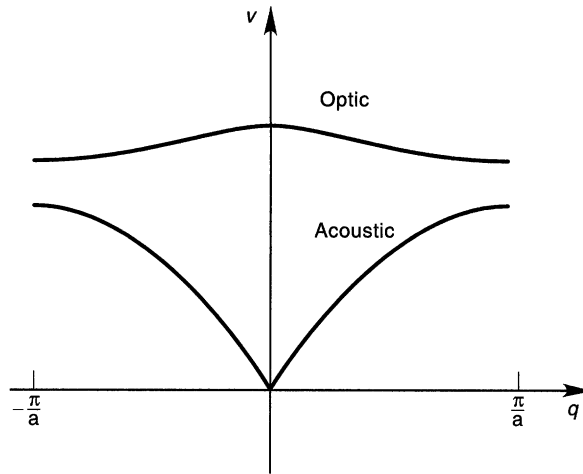


Figure 1.6 Schematic $v(q)$ diagram for acoustic and optic phonons in one dimension.

The energy versus q relation can then have two branches, the acoustic and the optic, as shown in Figure 1.6.

The presence of two different atoms can also cause long-range coulombic forces owing to the different charge on the two atom types (ionic component). The long-range forces cannot be described by simple forces between neighboring atoms, and one calls the phonons *polar optical phonons* if these long-range forces are important.

As mentioned, lattice vibrations are important in various ways. Electrons interact with the crystal lattice exciting (emitting) and absorbing lattice vibrations (the net lost energy is known as Joules heat). The system of electrons by itself is therefore not a Hamiltonian system; that is, one in which energy is conserved. It is only the sum of electrons and lattice vibrations which is Hamiltonian.

The interested reader is encouraged to obtain knowledge of a detailed quantum picture of lattice vibrations (also phonons) and their interactions with electrons as described, for example, by Landsberg [5].

1.2 THE EQUATIONS OF QUANTUM MECHANICS

At the beginning of the twentieth century, scientists realized that nature cannot be strictly divided into waves and particles. They found that light has particle-like properties and cannot always be viewed as a wave, and particles such as electrons revealed definite wave-like behavior under certain circumstances. They are, for example, diffracted by gratings as if they had a wavelength

$$\lambda = \frac{h}{|\mathbf{p}|} \quad (1.24)$$

where $\hbar \equiv h/2\pi \approx 6.58 \times 10^{-16}$ eVs is Planck's constant and $\mathbf{p}$ is the electron momentum.

Schrödinger demonstrated that the mechanics of atoms can be understood as boundary value problems. In his theory, electrons are represented by a wave function $\psi(\mathbf{r})$, which can have real and imaginary parts, and follows an eigenvalue differential equation:

$$\left(-\frac{\hbar^2}{2m} \nabla^2 + V(r) \right) \psi(r) = E\psi(r) \quad (1.25)$$

The part of the left side of Eq. (1.25) that operates on ψ is now called the Hamiltonian operator H . Formally this operator is obtained from the classical Hamiltonian by replacing momentum with the operator $\nabla\hbar/i$ (i = imaginary unit), where

$$\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right)$$

The meaning of the wave function $\psi(\mathbf{r})$ was not clearly understood at the time Schrödinger derived his famous equation. It is now agreed that $|\psi(\mathbf{r})|^2$ is the probability of finding an electron in a volume element $d\mathbf{r}$ at $\mathbf{r}$. In other words, we have to think of the electron as a point charge with a statistical interpretation of its whereabouts (the wave-like nature). It is usually difficult to get a deeper understanding of this viewpoint of nature; even Einstein had trouble with it. It is, however, a very successful viewpoint that describes exactly all phenomena we are interested in. To obtain a better feeling for the significance of $\psi(\mathbf{r})$, we will solve Eq. (1.25) for several special cases. As in the classical case, the simplest solution is obtained for constant potential. Choosing an appropriate energy scale, we put $V(\mathbf{r}) = 0$ everywhere.

By inspection we can see that the function

$$C \exp(i\mathbf{k} \cdot \mathbf{r}) = C(\cos \mathbf{k} \cdot \mathbf{r} + i \sin \mathbf{k} \cdot \mathbf{r}) \quad (1.26)$$

is a solution of Eq. (1.25) with

$$\frac{\hbar^2 k^2}{2m} = E \quad (1.27)$$

and C a constant.

The significance of the vector $\mathbf{k}$ can be understood from analogies to well-known wave phenomena in optics and from the classical equations. Because E is the kinetic energy, $\hbar\mathbf{k}$ has to be equal to the classical momentum $\mathbf{p}$ to satisfy $E = p^2/2m$. On the other hand, in optics

$$|\mathbf{k}| = 2\pi/\lambda \quad (1.28)$$

which gives, together with Eq. (1.24),

$$\hbar\mathbf{k} = \mathbf{p}$$

which is consistent with the mechanical result.

How can the result of Eq. (1.26) be understood in terms of the statistical interpretation of $\psi(\mathbf{r})$? Apparently

$$|\psi(\mathbf{r})|^2 = |C|^2(\cos^2 \mathbf{k} \cdot \mathbf{r} + \sin^2 \mathbf{k} \cdot \mathbf{r}) = |C|^2$$

This means that the probability of finding the electron at any place is equal to C^2 . If we know that the electron has to be in a certain volume V_{ol} (e.g., of a crystal), then the probability of finding the electron in the crystal must be one. Therefore,

$$\int_{V_{\text{ol}}} |C|^2 d\mathbf{r} = V_{\text{ol}} |C|^2 = 1$$

and

$$|C| = 1/\sqrt{V_{\text{ol}}} \quad (1.29)$$

In other words, the probability of finding an electron with momentum $\hbar\mathbf{k}$ at a certain point $\mathbf{r}$ is the same in the whole volume and equals $1/V_{\text{ol}}$. We will give a more detailed discussion of this somewhat peculiar result in the next section. The unfamiliar reader is referred to an introductory text (e.g., Feynman [3]).

Note that by confining the electron to a volume, we have already contradicted our assumption of constant potential $V(\mathbf{r}) = 0$. (Electrons can only be confined in potential wells.) If, however, the volume is large, our mistake is insignificant for many purposes.

Let us now consider the confinement of an electron in a one-dimensional potential well (although such a thing does not exist in nature). We assume that the potential energy $V(\mathbf{r})$ is zero over the distance $(0, L)$ on the x -axis and infinite at the boundaries 0 and L .

The Schrödinger equation, Eq. (1.25), reads in one dimension (x -direction, $V(x) = 0$)

$$-\frac{\hbar^2}{2m} \frac{\partial^2 \psi(x)}{\partial x^2} = E\psi(x) \quad (1.30)$$

Inspection shows that the function

$$\psi(x) = \sqrt{\frac{2}{L}} \sin \frac{n\pi}{L} x \quad \text{with } n = 1, 2, 3, \dots \quad (1.31)$$

satisfies Eq. (1.30) as well as the boundary conditions. The boundary conditions are, of course, that ψ vanishes outside the walls, since we assumed an infinite impenetrable potential barrier. In the case of a finite potential well, the wave function penetrates into the boundary and the solution is more complicated. If the barrier has a finite width, the electron can even leak out of the well (tunnel). This is a very important quantum phenomenon the reader should be familiar with. We will return to the tunneling effect below.

The wave function, Eq. (1.31), corresponds to energies E (called *eigen energies*)

$$E = \frac{n^2 \pi^2 \hbar^2}{2mL^2} \quad (1.32)$$

Because n is an integer, the electron can assume only certain discrete energies while other energies are not allowed. These discrete energies that can be assumed are called *quantum states* and are characterized by the *quantum number* n . The wave function and corresponding energy are therefore also denoted by ψ_n, E_n .

Think of a violin string vibrating in various modes at higher and lower tones (frequency ν), depending on the length L , and consider Einstein's law:

$$E = h\nu \quad (1.33)$$

If we compare the modes of vibration of the string with the form of the wave function for various n , then we can appreciate the title of Schrödinger's paper, "Quantization as a Boundary Value Problem."

Devices that contain a well and feature quantized energy levels similar to the ones given in Eq. (1.32) do exist. Quantum well lasers typically contain one or more small wells and the well size controls the electron energy. However, in most devices, the wells are not rectangular. In silicon metal oxide semiconductor field effect transistors (MOSFETs), the well is closer to triangular and its shape depends on the electron density (i.e., the charge in the well). We will deal with this charge-dependent well shape in Chapter 10. Here we discuss only well-defined potential problems—cases where the potential is given and fixed. With current high-end workstations, the Schrödinger equation can then be solved numerically for an arbitrary (but given) potential shape. One-dimensional problems can be solved by standard discretization (transforming the differentiations into finite differences) and by solving the resulting matrix equations by standard solvers such as found in EISPACK and LAPACK. For two- and three-dimensional problems, this procedure still leads to a prohibitively large number of equations (which grows with the third power of the number of discretization points). Therefore the discretized mesh must be coarsened even when using the fastest supercomputers. Often, however, one is interested only in relatively small sets of eigenvalues, for example, the first three [as for $n = 1, 2, 3$ in Eq. (1.31)].

One then can use so-called subspace interaction techniques that only resolve certain intervals of eigenvalues. These techniques are well established for symmetric real matrices as they occur in well-defined potential problems (see, e.g., Golub and Loan [4]). A useful computer code is the RITZED eigenvalue solver by Rutishauser [6].

Frequently, one needs to obtain an explicit solution of Schrödinger's equation for an arbitrary complicated form of the potential, provided only that it is small and represents just a small perturbation to a problem for which the solution is known. This scenario is typical for scattering problems such as an electron propagating in a perfect solid and then encountering a small imperfection and being scattered. Fortunately, for this type of problem there is a powerful method of approximation, perturbation theory, that gives us the solution for arbitrary weak potentials. The method is very general and applies to any kind of equation.

Consider an equation of the form

$$(H_0 + \varepsilon H_1)\psi = 0 \quad (1.34)$$

where H_0 and H_1 are differential operators of arbitrary complication and ε is a small positive number.

If we know the solution ψ_0 of the equation

$$H_0\psi_0 = 0$$

then we can assume that the solution of Eq. (1.34) has the form $\psi_0 + \varepsilon\psi_1$. Inserting this form into Eq. (1.34), we obtain

$$(H_0 + \varepsilon H_1)(\psi_0 + \varepsilon\psi_1) = H_0\psi_0 + \varepsilon H_1\psi_0 + H_0\varepsilon\psi_1 + \varepsilon^2 H_1\psi_1$$

We now can neglect the term proportional to ε^2 (because ε is small), and because $H_0\psi_0 = 0$, we have

$$H_1\psi_0 + H_0\psi_1 = 0 \quad (1.35)$$

This equation is now considerably simpler than Eq. (1.34) because ψ_0 is known. Therefore, ψ_1 can be determined easily if H_0 has a simple form no matter how complicated H_1 is. Repeated application of this principle leads to perturbation theory including higher orders ($\varepsilon_1^2\varepsilon^3\dots$). The derivation is given in many textbooks on quantum mechanics (see Baym [1]). Here we quote only the result that is used at several occasions.

Assume that we know the solutions of a Schrödinger equation:

$$H_0\psi_n = E_n\psi_n \quad n = 1, 2, 3, \dots \quad (1.36)$$

and we would like to know the solutions of

$$(H_0 + H_1)\phi_m = W_m\phi_m \quad \text{with} \quad H_1 \ll H_0 \quad (1.37)$$

Then it is shown in elementary texts on quantum mechanics (Baym [1]), by repeatedly using the method of perturbation theory as outlined, that

$$W_m = E_m + M_{mm} + \sum_{n \neq m} \frac{|M_{mn}|^2}{E_m - E_n} \quad (1.38)$$

with

$$\phi_m = \psi_m + \sum_{n \neq m} \frac{M_{mn}}{E_m - E_n} \psi_n \quad (1.39)$$

and

$$M_{mn} = \int_{V_{\text{ol}}} \psi_n^* H_1 \psi_m d\mathbf{r} \quad (1.40)$$

where $d\mathbf{r}$ stands for $dx dy dz$ (integration over volume V_{ol}) and ψ_n^* is the complex conjugate of ψ_n .

First-order perturbation theory (to order ε) amounts to setting $\psi_m = \Phi_m$ and $W_m = E_m + M_{mm}$. The only change then is in the value of the eigen energy by M_{mm} , which can be obtained by the integration in Eq. (1.40); the integrand is known from the solution of Eq. (1.36). This means that the numerical problem

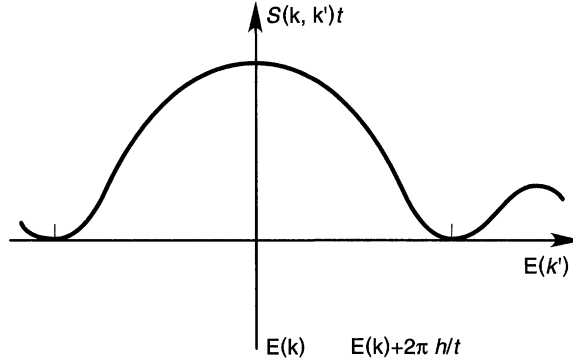


Figure 1.7 Probability of a transition from $\mathbf{k}$ to $\mathbf{k}'$ according to the Golden Rule after the potential has been on for time t .

is reduced to a volume integration (in three dimensions). To obtain solutions to higher order, one also needs to perform the summations in Eqs. (1.38) and (1.39).

The formalism outlined above and the examples given are independent of time, and the electrons are perpetually in appropriate (eigen) states. In many instances, however, we will be interested in the following type of problem: The electron initially is in an eigenstate of H_0 , denoted, for example, by a wave vector $\mathbf{k}$ for the free electron. What is the probability that the electron will be observed in a different eigenstate characterized by the wave vector $\mathbf{k}'$, after it interacts with a potential $V(\mathbf{r}, t)$? In other words, what is the probability $S(\mathbf{k}, \mathbf{k}')$ per unit time that the interaction causes the system to make a transition from $\mathbf{k}$ to $\mathbf{k}'$?

The answer to this question is the famous Golden Rule of Fermi, which is also derived in almost every text on quantum mechanics by so-called time-dependent perturbation theory. The unfamiliar reader is urged to acquire a detailed understanding of the Golden Rule as derived, for example, in the text of Baym [1]. Here we only illustrate its generality and discuss results for important special cases.

1. Assume that a potential $V(\mathbf{r})$ is switched on at time $t = 0$ but is time independent otherwise. One then obtains

$$S(\mathbf{k}, \mathbf{k}') = \left| \int_{V_{\text{ol}}} \psi_{\mathbf{k}'}^* V(\mathbf{r}) \psi_{\mathbf{k}} d\mathbf{r} \right|^2 \cdot \left[\frac{\sin(E(\mathbf{k}') - E(\mathbf{k}))t/2\hbar}{(E(\mathbf{k}') - E(\mathbf{k}))\sqrt{t/2}} \right]^2 \quad (1.41)$$

The function in brackets deserves special attention and is plotted in Figure 1.7. Notice that as t approaches infinity, the function plotted in Figure 1.7 becomes more and more peaked at its center ($E(\mathbf{k}') = E(\mathbf{k})$). In the limit $t \rightarrow \infty$, the so-called δ -function is approached, which is defined by

$$\lim_{t \rightarrow \infty} \frac{4 \sin^2[(E(\mathbf{k}') - E(\mathbf{k}))t/(2\hbar)]}{(E(\mathbf{k}') - E(\mathbf{k}))^2 t} = \frac{2\pi}{\hbar} \delta(E(\mathbf{k}') - E(\mathbf{k})) \quad (1.42)$$

and can always be understood as a limit of ordinary functions. It does have some remarkable properties, however, and the unfamiliar reader should consult some of the references at the end of this section. A most important property of the δ -function is the following: For any continuous function $f(E')$, we have

$$\int_{-\infty}^{\infty} f(E') \delta(E - E') dE' = f(E) \quad (1.43)$$

2. We assume that the perturbation is harmonic, which means we have a potential of the form

$$V(\mathbf{r}, t) = V(\mathbf{r})(e^{-i\omega t} + e^{i\omega t})$$

For $t \rightarrow \infty$, we obtain the transition probability

$$S(\mathbf{k}, \mathbf{k}') = \frac{2\pi}{\hbar} \left| \int_{V_{\text{ol}}} \psi_{\mathbf{k}'}^* V(\mathbf{r}) \psi_{\mathbf{k}} d\mathbf{r} \right|^2 \left[\delta(E(\mathbf{k}) - E(\mathbf{k}') - \hbar\omega) + \delta(E(\mathbf{k}) - E(\mathbf{k}') + \hbar\omega) \right] \quad (1.44)$$

It is clear that the δ -function simply takes care of energy conservation. For a constant potential, we have to conserve energy as t increases. For a harmonic perturbation, the system can gain or lose energy corresponding, for example, to the absorption or emission of light.

3. We now turn our attention to the first term in Eq. (1.41), the matrix element, which also plays a vital role in time-independent perturbation theory. The significance of the matrix element is best illustrated by the following special cases of well-defined potential problems (problems in which the potential is given by a certain function of coordinates):

- (a) $V(\mathbf{r}) = \text{constant}$. The matrix element is then

$$\text{constant} \frac{1}{V_{\text{ol}}} \int_{V_{\text{ol}}} e^{-i\mathbf{k}' \cdot \mathbf{r}} e^{i\mathbf{k} \cdot \mathbf{r}} d\mathbf{r} \quad (1.45)$$

The integration is over the volume V_{ol} of the crystal. In many practical cases, this volume will be much larger than the de Broglie wavelength λ of the electron, which is of the order of 100 Å in typical semiconductor problems. This means that the integral of Eq. (1.45) will be very close to zero, because the cosine and sine functions to which the exponents in Eq. (1.45) are equivalent are positive as often as they are negative in the big volume. There is only one exception: In the case $\mathbf{k}' = \mathbf{k}$, the integral is equal to the volume and the matrix element is equal to constant. Therefore, we can write

$$\text{constant} \frac{1}{V_{\text{ol}}} \int_{V_{\text{ol}}} e^{-i\mathbf{k}' \cdot \mathbf{r}} e^{i\mathbf{k} \cdot \mathbf{r}} d\mathbf{r} = \text{constant} \delta_{\mathbf{k}', \mathbf{k}} \quad (1.46)$$

where $\delta_{\mathbf{k}', \mathbf{k}} = 1$ for $\mathbf{k} = \mathbf{k}'$ and is zero otherwise. This is known as the *Kronecker delta symbol*. Consequently, the matrix element has

taken care of momentum conservation; the free electron in a constant potential does not change its momentum.

- (b) Second, we consider an arbitrary potential having the following Fourier representation:

$$V(\mathbf{r}) = \sum_{\mathbf{q}} V_{\mathbf{q}} e^{i\mathbf{q} \cdot \mathbf{r}} \quad (1.47)$$

Then the matrix element, which we now denote by $M_{\mathbf{k},\mathbf{k}'}$, becomes

$$\begin{aligned} M_{\mathbf{k},\mathbf{k}'} &= \sum_{\mathbf{q}} \frac{V_{\mathbf{q}}}{V_{\text{ol}}} \int_{V_{\text{ol}}} e^{i(\mathbf{k}-\mathbf{k}'+\mathbf{q}) \cdot \mathbf{r}} \\ &= \sum_{\mathbf{q}} V_{\mathbf{q}} \delta_{\mathbf{k}'-\mathbf{k},\mathbf{q}} \end{aligned} \quad (1.48)$$

How do we interpret this result? If we also allow the potential of Eq. (1.47) to have a time dependence (e.g., as $e^{i\omega t}$), then the potential can be interpreted as that of a wave (e.g., an electromagnetic wave). In this case Eq. (1.49) simply tells us that the wave vectors of all scattering agents (i.e., their momenta) are conserved, because we have

$$\mathbf{k}' - \mathbf{k} = \mathbf{q} \quad (1.49)$$

It is important to notice that Eq. (1.49) is also valid for a static, time-independent potential—that is, even a static potential “supplies” momentum according to its Fourier components—in the same way a wave does. This seems strange at first glance. To see the significance, consider the boundary of a billiard table. This boundary is an impenetrable abrupt potential step whose Fourier decomposition involves all values of $\mathbf{q}$. Indeed, the boundary can supply any momentum to the ball to make it bounce back. The above two examples show that the Golden Rule essentially takes care of energy and momentum conservation. This is also the reason for its generality and importance. Remember, however, that this is true only for cases when time t at which we observe the scattered particle is long after the potential is switched on. For short times (in practice these are times of the order of 10^{-14} s), the function in Eq. (1.42) cannot be approximated by a δ function, and energy need not be conserved in processes on this short time scale. This is at the heart of the energy time uncertainty relation. To illustrate the great generality of the Golden Rule, one more example is given.

- (c) Consider the “tunneling problem” of Figure 1.8. Although an electric field F is applied in the z -direction, the electron in Figure 1.8 is confined in a small well. Classically, it would stay in the well. However, because the barrier is not infinite, as assumed in Eq. (1.31), the wave function is not zero at the well boundary but penetrates the

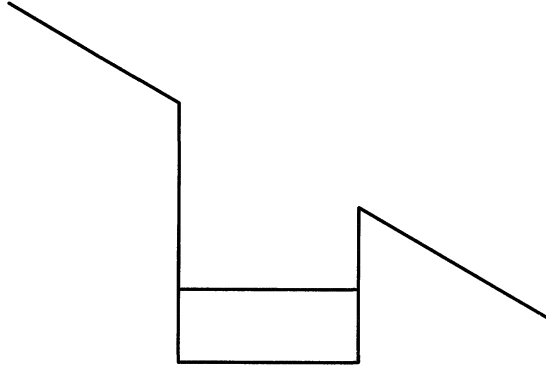


Figure 1.8 Electrons in a potential well plus applied electric field F .

boundary. In other words, there is a finite probability of finding the electron outside the well.

We can calculate the probability per unit time that the electron leaks out if we know the wave function ψ_{in} in the well and ψ_{ou} outside the well, and regard the electric field as a perturbation. This perturbation gives a term (the potential energy) eFz in the Hamiltonian. The Golden Rule tells us that

$$S(w, ou) = \left| \int_{V_{ol}} \psi_{ou}^* eFz \psi_{in} d\mathbf{r} \right| \cdot \frac{2\pi}{\hbar} \delta(E_{ou} - E_{in}) \quad (1.50)$$

In writing down this equation (which was first derived by Oppenheimer), I have swept under the rug the fact that ψ_{ou} and ψ_{in} are the solutions of different Hamiltonians. ψ_{in} is obtained from the solution of the Schrödinger equation of the quantum well and ψ_{ou} is the solution of a free electron in an electric field with

$$H = -\frac{\hbar^2 \nabla^2}{2m} - eFz$$

An exact justification of this procedure is complicated and is discussed in great detail in Duke's treatise of tunneling [2] (see also Appendix A).

We emphasize that the matrix elements represent all that needs to be known to obtain perturbation theory solutions. These matrix elements are given by three dimensional integrals. Alternatively they can be viewed as scalar products in a vector space denoting ψ_n by a vector $|n\rangle$. For those unfamiliar with Dirac's notation the following definition can just be used as a shorthand way of writing the integral:

$$\langle m | H_1 | n \rangle = \int_{V_{ol}} \psi_m^* H_1 \psi_n d\mathbf{r} \quad (1.51)$$

PROBLEMS

- 1.1 Solve by perturbation theory (to first order) for y :

$$\varepsilon \frac{\partial y \sin(x)}{\partial x} + \frac{\partial y}{\partial x} = y$$

where ε is a small positive quantity.

- 1.2 Calculate the matrix elements for wave functions of the form $\psi = \frac{1}{\sqrt{V_{\text{ol}}}} e^{i\mathbf{k} \cdot \mathbf{r}}$ and

(a) $H_1 \propto e^{i\mathbf{q} \cdot \mathbf{r}}$

(b) $H_1 \propto \delta(\mathbf{r})$

(c) $H_1 \propto |\mathbf{r}|^{-2}$

(d) $H_1 \propto \exp\left\{\frac{-|\mathbf{r}|}{r_0}\right\} \quad r_0 > 0$ Polar coordinates are helpful in parts c and d.

- 1.3 Consider a one-dimensional crystal lattice with two ions (atoms) repeated in a circular arrangement. The two ions (atoms) are identical, with mass M , but are connected by springs of alternating strength (D_1, D_2) .

(a) Derive the equations of motion. (Consider only nearest neighbor interactions, where the force is proportional to the difference in displacements.)

(b) Find and sketch the dispersion relation of the possible vibrational modes. (Assume all displacements are traveling waves with sinusoidal time dependence, that is, $u_i(ra) = \varepsilon_i e^{i(qra - \omega t)}$.)

(c) Discuss the form of the dispersion relation and the nature of the modes for $q \ll \pi/a$ and $q = \pi/a$, where q is the wave vector.

(d) Find the velocity of sound (ω/q for $q \rightarrow 0$).

(e) Show that the group velocity $\partial\omega/\partial q$ becomes zero at the Brillouin zone boundary. (This is a general result.)

REFERENCES

- [1] Baym, G. *Lectures on Quantum Mechanics*. New York: Benjamin, 1969, p. 248.
- [2] Duke, C. B. *Tunneling in Solids*. New York: Academic Press, 1969, p. 207.
- [3] Feynman, R. P. *Lectures on Physics*, Vol. III. Reading, MA: Addison-Wesley, 1964, pp. 1.1–4.15.
- [4] Golub, G. H., and Loan, C. F. *Matrix Computation*. John Hopkins Univ. Press: Baltimore, 1989.
- [5] Landsberg, P. T. *Solid State Theory*. New York: Wiley/Interscience, 1969, pp. 327–56.
- [6] Rutishauser, H. “Numerical eigenvalue solver,” *Numerical Mathematics*, vol. 13, 1969, p. 4.