Chapter 12

Time-Independent Perturbation Theory

12.1   Introduction

In chapter 3 we discussed a few *exactly* solved problems in quantum mechanics. Many applied problems may not be exactly solvable. The machinery to solve such problems is called perturbation theory. In chapter 11, we developed the matrix formalism of quantum mechanics, which is well-suited to handle perturbation theory. Sometimes we will be able to reduce the matrix solutions to closed-form algebraic forms which always helps in visualization.

Let $\hat{H}^0$ be the Hamiltonian for the solved problem. Then the time-dependent Schrodinger equation is $i\hbar\frac{\partial}{\partial t}|\Psi\rangle = \hat{H}^0|\Psi\rangle$. The eigenstates of definite energy are also *stationary* states $\langle r|n\rangle = \psi_E(r)e^{-iE_n t/\hbar}$, where $|n\rangle$ are the eigenvectors and E_n the corresponding eigenvalues. Note that *all* the solved problems we discussed in chapter 3 such as the harmonic oscillator or the particle in a box had *time-independent* potentials. Many real-world situations involve time-dependent potentials. For example, imagine a field-effect transistor whose gate voltage is being modulated by an ac signal. That will create a potential variation for electrons of the form $V(r)e^{i\omega t}$. A similar variation will be experienced by electrons interacting with photons of an electromagnetic wave, or with phonons of lattice vibrations. Consider the limit of very small frequencies $\omega \to 0$, or a 'dc' potential. Then, the potential only has a spatial variation. A dc voltage is not truly time-independent because it has to be turned on or off at some time. But most of the physics we are interested in this and a few following chapters happens when the perturbation is 'on', and things have reached steady-state. It is in this sense that we discuss time-independent

perturbation theory. We defer explicitly time-varying or oscillatory perturbations to chapter 22.

12.2 Degenerate Perturbation Theory

The time-independent Schrodinger equation for the *solved* problem is

$$\hat{H}^0|n\rangle = E_n^0|n\rangle, \tag{12.1}$$

where $\hat{H}^0$ is the unperturbed Hamiltonian. That means we know all the eigenfunctions $|n\rangle$ and their corresponding eigenvalues E_n^0. This is shown in Fig 12.1. Lets add a perturbation W to the initial Hamiltonian such that the new Hamiltonian becomes $\hat{H} = \hat{H}^0 + W$. The new Schrodinger equation is

$$(\hat{H}^0 + W)|\psi\rangle = E|\psi\rangle. \tag{12.2}$$

The perturbation W has changed the eigenvectors $|n\rangle \to |\psi\rangle$. The corresponding eigenvalues may not be eigenvalues of the new Hamiltonian. Some eigenvalues increase in energy, some decrease, and others may not be affected. This is illustrated in Fig 12.1. So we have to solve for the new eigenvalues E and obtain the corresponding eigenvectors.
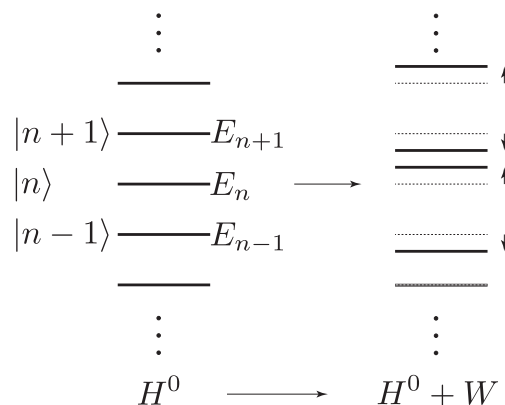

FIGURE 12.1: The initial eigenstates and eigenvalues of a quantum system change upon application of a perturbation W.

At this stage, we invoke the *Expansion Principle* introduced in chapter 11. It states that the *perturbed* state vector $|\psi\rangle$ can always be written as a linear superposition of the *unperturbed* eigenvector $|n\rangle$, since the unperturbed eigenstates formed a complete basis. It is the same philosophy of expanding any function in terms of its Fourier components. Thus we write

$$|\psi\rangle = \sum_n a_n |n\rangle, \tag{12.3}$$

where a_n's are (in general complex) expansion coefficients. The coefficients are obtained by taking the projection $\langle m|\psi\rangle$, which yields $a_n = \langle n|\psi\rangle$. Then equation 12.2 reads

$$\sum_n a_n(\hat{H}^0 + W)|n\rangle = E\sum_n a_n|n\rangle. \tag{12.4}$$

We can visualize the new state vector as the original eigenvector 'rotated' by the perturbation W. Lets project the new state vector on $\langle m|$ to get

$$\sum_n a_n\langle m|(\hat{H}^0 + W)|n\rangle = Ea_m, \tag{12.5}$$

which is a matrix when m takes values $1, 2, \ldots N$

$$\begin{bmatrix} E_1 + W_{11} & W_{12} & W_{13} & \ldots & W_{1N} \\ W_{21} & E_2 + W_{22} & W_{23} & \ldots & W_{2N} \\ W_{31} & W_{32} & E_3 + W_{33} & \ldots & W_{3N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ W_{N1} & W_{N2} & W_{N3} & \ldots & E_N + W_{NN} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_N \end{bmatrix} = E \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_N \end{bmatrix}. \tag{12.6}$$

The eigenvalues and the corresponding eigenvectors of this matrix equation are obtained by diagonalization, as discussed in chapter 11. The new eigenvalues E_n' thus depend on the matrix elements $W_{mn} = \langle m|W|n\rangle$ of the perturbation. Note that if some eigenvalues of the unperturbed Hamiltonian happened to be degenerate, the matrix diagonalization method takes that into account naturally without problems. In that sense, the matrix formulation of perturbation theory is sometimes referred to as *degenerate* perturbation theory. But the matrix formulation handles non-degenerate situations equally well, and is more general.

In case we did not start with a diagonal basis of the unperturbed Hamiltonian H^0, then we have the Schrodinger equation

$$\begin{bmatrix} H_{11}^0 + W_{11} & H_{12}^0 + W_{12} & H_{13}^0 + W_{13} & \ldots & H_{1N}^0 + W_{1N} \\ H_{21}^0 + W_{21} & H_{22}^0 + W_{22} & H_{23}^0 + W_{23} & \ldots & H_{2N}^0 + W_{2N} \\ H_{31}^0 + W_{31} & H_{32}^0 + W_{32} & H_{33}^0 + W_{33} & \ldots & H_{3N}^0 + W_{3N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ H_{N1}^0 + W_{N1} & H_{N2}^0 + W_{N2} & H_{N3}^0 + W_{N3} & \ldots & H_{NN}^0 + W_{NN} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_N \end{bmatrix} = E \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_N \end{bmatrix}. \tag{12.7}$$

The solutions thus reduce to diagonalizing the corresponding perturbed matrices. Many of the perturbation matrix elements $W_{mn} = \langle m|W|n \rangle$ can be made zero by suitable choice of bases, which reduces the work involved in diagonalization. Note that we can easily obtain the eigenvalues, the eigenvectors typically require more work. But with the availability of math packages such as Mathematica and MATLAB, this is done in a jiffy for most situations we will deal with.

An important question in applying degenerate perturbation theory is: which states should be included in the $N \times N$ matrix? At this stage, we state the guiding principles, we will see the proof of the principle in the next section on non-degenerate perturbation theory. The first principle is that states with eigenvalues widely separated in energy $E_n - E_m$ interact weakly by the perturbation. The second principle is states get pushed around by the perturbation depending upon the matrix element squared $|W_{mn}|^2$. Quantitatively, the perturbation in energy of a state with energy E is $\Delta E \approx |W_{mn}|^2/(E-E_n)$ by interacting with state $|n\rangle$. Thus, states for which W_{mn} terms are small or zero may be left out. If we are interested in a set of energy eigenvalues (say near the conduction and valence band edges of a semiconductor), energies far away from the band edges may also be left out. We will see the application of these rules in chapters 13 and many following.

We will see examples of degenerate perturbation theory in the next chapter (chapter 13), where we will apply it to the problem of a free electron. That will require us to solve either 2×2 matrices, or higher order, depending on the accuracy we need. Later on, we will also encounter it when we discuss the $\mathbf{k} \cdot \mathbf{p}$ theory of bandstructure to deal with the degeneracies of heavy and light valence bands. For now, we look at particular situations when we have 'isolated' eigenvalues that are non-degenerate.

12.3   Non-Degenerate Perturbation Theory

Schrodinger's crowning achievement was to obtain an algebraic equation, which when solved, yields the quantum states allowed for electrons. Schrodinger's equation is called

the 'wave'-equation because it was constructed in analogy to Maxwell's equations for electromagnetic waves. Heisenberg was the first to achieve the breakthrough in quantum mechanics before Schrodinger, except his version involved matrices. Which is why he called it matrix-mechanics. That is why it was not as readily accepted - again because matrices are unfamiliar to most. It was later that the mathematician von Neumann proved that both approaches were actually identical from a mathematical point of view.

So at this point, we will try to return to a 'familiar' territory in perturbation theory from the matrix version presented in the previous section. We try to formulate an *algebraic* method to find the perturbed eigenvalues and eigenvectors.
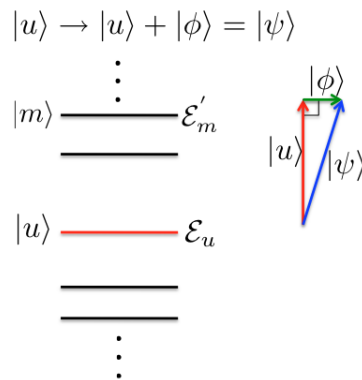

FIGURE 12.2: The perturbation rotates the eigenvector $|u\rangle$ to $|\psi\rangle$. If we forego normalization of $|\psi\rangle$, we can find a vector $|\phi\rangle$ orthogonal to $|u\rangle$ such that $\langle u|\phi\rangle = 0$, and consequently $\langle u|\psi\rangle = 1$.

Consider a perturbation $\hat{W}$ added to the solved (or unperturbed) Hamiltonian $\hat{H}^0$. Schrodinger equation is

$$(\hat{H}^0 + \hat{W}))|\psi\rangle = E|\psi\rangle, \tag{12.8}$$

and the unperturbed state $|u\rangle$ satisfied

$$\hat{H}^0|u\rangle = E_u|u\rangle. \tag{12.9}$$

The new quantum state differs from the unperturbed state, so we write

$$|\psi\rangle = |u\rangle + |\phi\rangle. \tag{12.10}$$

We can picturize the final state $|\psi\rangle$ as a 'vector' sum of the unperturbed state $|u\rangle$ and the vector $|\phi\rangle$. This is schematically shown in Fig 12.2. In particular, if we are willing to not normalize the final state, then we can always choose $|\phi\rangle$ to be orthogonal to $|u\rangle$,

leading to $\langle u|\phi\rangle = 0$ and $\langle u|\psi\rangle = 1$. We can then project equation 12.8 on $\langle u|$ to obtain the energy equation

$$E = E_u + \langle u|W|\psi\rangle = \underbrace{E_u}_{\text{unperturbed}} + \underbrace{\langle u|W|u\rangle}_{\Delta E^{(1)}} + \underbrace{\langle u|W|\phi\rangle}_{\text{higher orders}} . \tag{12.11}$$

Note that we obtain the 'first-order' energy correction: they are the diagonal matrix elements of the perturbation with the unperturbed states. Think of a 'dc' perturbation - say a voltage V_0 that depends neither on space nor time - then all the initial energy eigenvalues get shifted by the corresponding energy: $E_u \to E_u + qV_0$ due to the first order term since $\langle u|qV_0|u\rangle = qV_0$. We will shortly see that for this particular perturbation, the higher order terms are zero because they depend on the *cross*-matrix terms of the kind $\langle m|qV_0|n\rangle = qV_0\langle m|n\rangle = 0$. An example of such a situation is when a voltage is applied across a gate capacitor to a semiconductor - the entire bandstructure which consists of the allowed E_n's shift rigidly up or down. We call this energy band-bending in device physics. Such a 'dc' perturbation does not couple different energy states for the above reason, and results in only a first-order rigid shift.

Most perturbations are not the 'dc'-kind, and we need the higher order terms for them. To do that, it is useful to define

$$E'_u = E_u + \langle u|W|u\rangle. \tag{12.12}$$

We then split the diagonal and off-diagonal elements of the perturbation just like writing a signal as a 'dc' + a 'ac' terms. Think of $\hat{W}$ as an operator, and hence a matrix that we are splitting it into two:

$$\hat{W} = \hat{D} + \hat{W}', \tag{12.13}$$

and the total Hamiltonian then is

$$\hat{H} = \underbrace{\hat{H}^0 + \hat{D}}_{\hat{H}^{(d)}} + \hat{W}', \tag{12.14}$$

The reason for doing this is that the unperturbed eigenvalues are going to shift by the diagonal part of the perturbation without interacting with other states. The off-diagonal terms will further tweak them by interactions with other states. To move further, we write

$$(\hat{H}^{(d)} + \hat{W}')|\psi\rangle = E|\psi\rangle, \tag{12.15}$$

and rearrange it to

$$(E - \hat{H}^{(d)})|\psi\rangle = \hat{W}'|\psi\rangle, \tag{12.16}$$

At this stage, our goal is to find the perturbation vector $|\phi\rangle = |\psi\rangle - |u\rangle$. How can we obtain it from the left side of equation 12.16 in terms of the perturbation on the right? Recall in chapter 11 we discussed the Green's function operator briefly. We noticed that it is an 'inverse' operator, meaning we expect

$$\hat{G}(E)(E - \hat{H}^{(d)})|\psi\rangle = \sum_m \frac{|m\rangle\langle m|}{E - E'_m}(E - \hat{H}^{(d)})|\psi\rangle = \sum_m |m\rangle\langle m|\psi\rangle = |\psi\rangle. \tag{12.17}$$

So to get $|\phi\rangle = |\psi\rangle - |u\rangle$, perhaps we should use the operator

$$\hat{G}(E) - \frac{|u\rangle\langle u|}{E - E'_u} = \sum_{m \neq u} \frac{|m\rangle\langle m|}{E - E'_m}. \tag{12.18}$$

Operating on the LHS of equation 12.16 we obtain

$$\sum_{m \neq u} \frac{|m\rangle\langle m|}{E - E'_m}(E - \hat{H}^{(d)})|\psi\rangle = \sum_{m \neq u} |m\rangle\langle m|\psi\rangle = \left(\sum_m |m\rangle\langle m|\psi\rangle\right) - |u\rangle\langle u|\psi\rangle = |\psi\rangle - |u\rangle = |\phi\rangle, \tag{12.19}$$

which is what we wanted. Now we use the same operator on the right of equation 12.16 to finish the job. Since $\hat{W}'$ consists of only off-diagonal cross matrix elements, we write it in its outer product form as $\hat{W}' = \sum_m \sum_{m \neq n} |m\rangle\langle m|\hat{W}|n\rangle\langle n|$, and apply the 'reduced' Green's function to get

$$|\phi\rangle = \sum_{l \neq u} \sum_m \sum_{n \neq m} \frac{|l\rangle\langle l|}{E - E'_l}|m\rangle\langle m|\hat{W}|n\rangle\langle n|\psi\rangle = \sum_{m \neq u} \sum_{n \neq m} |m\rangle \frac{\langle m|\hat{W}|n\rangle}{E - E'_m}\langle n|\psi\rangle, \tag{12.20}$$

Thus, we obtain the perturbed state $|\psi\rangle = |u\rangle + |\phi\rangle$ to be

$$|\psi\rangle = |u\rangle + \underbrace{\sum_{m \neq u} \sum_{n \neq m} |m\rangle \frac{\langle m|\hat{W}|n\rangle}{E - E'_m} \langle n|\psi\rangle}_{|\phi\rangle} . \qquad (12.21)$$

As a sanity check, we note that if $\hat{W} = 0$, $|\psi\rangle = |u\rangle$, as it should be. Next, we note that this is a *recursive* relation, meaning $|\psi\rangle$ also appears inside the sum on the right side. Thus, it can be taken to many orders, but we are going to retain just up to the 2nd order. That means, we will assume that the perturbation is weak, and so we are justified in replacing the $|\psi\rangle$ inside the sum on the right side by the unperturbed state $|u\rangle$. With that, we get the result

$$\boxed{|\psi\rangle \approx |u\rangle + \underbrace{\sum_{m \neq u} \frac{\langle m|\hat{W}|u\rangle}{E - E'_m} |m\rangle}_{\phi^{(1)}} .} \qquad (12.22)$$

The perturbed state vector given by equation 12.22 now is in a way that can be used for calculations. That is because every term on the right side is known. To obtain the perturbed eigenvalues, we substitute equation 12.2 into the expression for energy $E = E_u + \langle u|W|u\rangle + \langle u|W|\phi\rangle$ to obtain

$$\boxed{E \approx E_u + \underbrace{\langle u|W|u\rangle}_{\Delta E^{(1)}} + \underbrace{\sum_{m \neq u} \frac{|\langle m|\hat{W}|u\rangle|^2}{E - E'_m}}_{\Delta E^{(2)}} .} \qquad (12.23)$$

This result is called the **Brillouin-Wigner** (BW) perturbation theory. Note that the BW algebraic solution for determining the unknown eigenvalues E require us to solving for it. But for multiple states, the solution would require a high order polynomial, since equation 12.23 is indeed a polynomial. For example, lets say we were looking at a 3-state problem with unperturbed energies E_{u1}, E_{u2}, E_{u3}, and we want to find how eigenvalues of state $u = 2$ got modified by the perturbation. Then, the 2nd energy energy correction has 2 terms, since $m \neq 2$. The equation then becomes a 3rd-order polynomial with three roots, which are the eigenvalues.

Such solving of polynomial equations can be avoided if we are willing to compromise on the accuracy. If so, the unknown energy term E in the denominator of the 2nd order correction term may be replaced by the unperturbed value, $E \to E_u$. Then the energy eigenvalues are obtained directly from

$$E \approx E_u + \langle u|\hat{W}|u\rangle + \sum_{m \neq u} \frac{|\langle m|\hat{W}|u\rangle|^2}{E_u - E'_m}. \tag{12.24}$$

this equation is called the **Rayleigh-Schrodinger** (RS) perturbation theory. Note that in this form, we know all the terms on the right side. It was first derived by Schrodinger right after his seminal work on the wave equation of electrons. The RS-theory is not applicable for understanding perturbation of degenerate states, as the denominator $E_n - E'_m$ can go to zero. But BW-theory applies for degenerate states too, and one can always resort back to the degenerate perturbation theory.
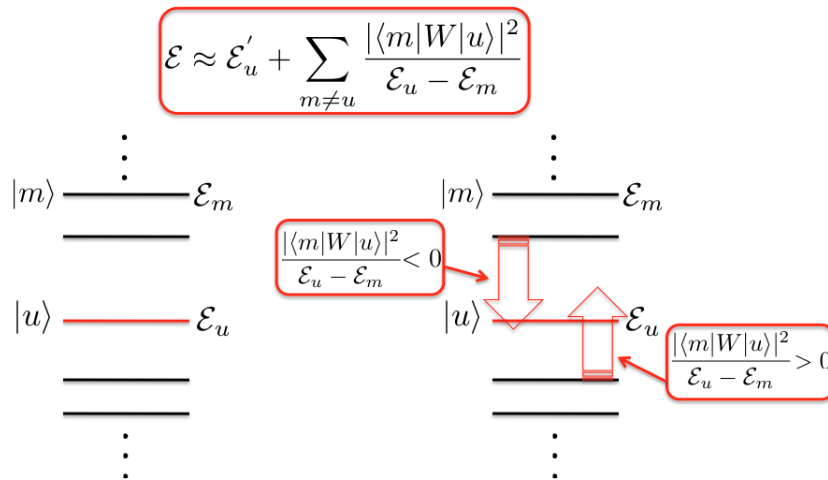

FIGURE 12.3: Illustration of revel repulsion due to perturbation.

In the treatment of degenerate perturbation theory earlier, we discussed the strategy to follow to choose which states to include in the matrix. The last term in the BW or RS perturbation theory results provides the guiding principle. Note that this term goes as the perturbation matrix element *squared*, divided by the energy difference. In the absence of the perturbation, the eigenvectors corresponding to the eigenvalues were orthogonal, meaning they did not 'talk' to each other. The perturbation mixes the states, and makes them talk. The magnitude by which the energy of a state $|u\rangle$ changes due to interactions with *all other states* upon perturbation is $\Delta E^{(2)} \approx \sum_{m \neq u} |W_{mu}|^2/(E_u - E'_m)$.

We also note the nature of the interaction. If a state E_u is interacting with states with energies E'_m lower than itself, then $\Delta E^{(2)} > 0$, the perturbation pushes the energy *up*. Similarly, interactions with states with higher energies pushes the energy of state E_u *down*. Thus, the second-order interaction term in perturbation is *repulsive*. Figure 12.3 illustrates this effect schematically. This repulsive interaction is the key to understanding curvatures of energy bands and the relation between effective masses and energy bandgaps of semiconductors. Clearly if two states were non-degenerate and the strength

of the perturbation is increased from zero, the energy eigenvalues repel stronger, and the levels go farther apart. Then they cannot cross each other. This is a case of what goes by the name of **no level crossing theorem** in perturbation theory.

In this chapter, we developed the theoretical formalism for handling time-independent perturbation theory. The matrix formalism is well-suited for uncovering the effect of perturbation on eigenvectors and eigenvalues. It works for problems where the unperturbed states are either degenerate, or non-degenerate energy states. For non-degenerate eigenstates, algebraic solutions can be obtained in the Brillouin-Wigner (BW), or the Rayleigh-Schrodinger (RS) theories. The analytical solutions offer further insights into the effect of the perturbation on the physical parameters of interest in the problems. In the next few chapters, we apply both degenerate and non-degenerate perturbation theories to understand electron bandstructure in semiconductors, and its various ramifications.

Debdeep Jena: www.nd.edu/~djena

Chapter 13

Free Electron Perturbed by a Periodic Potential

13.1   Introduction

In the last chapter, we developed the formalism for time-independent perturbation theory. In this chapter, we apply the theory one problem: that of a free electron perturbed by a periodic potential. The results we obtain will highlight most of the fundamental properties of semiconductors. These include their energy bandstructure and opening of bandgaps, evolution of effective masses of various bands, work function, interactions between electron states in solids, and the role of defects in the interactions between electron states. In brief, the chapter will capture the essence of time-independent semiconductor physics. Much of the following chapters are detailed treatments of increasing levels of sophistication, till we need time-dependent behavior which will require new concepts. The central time-independent phenomena are captured in this chapter.

13.2   The free-electron

In chapter 3, we discussed the free electron problem. Here we discuss purely the 1-D problem. For the free electron, the potential term in the Schrodinger equation is zero $V(x) = 0$. The eigenvectors $|k\rangle$ are such that their real-space projection yields the plane wave-function

$$\langle x|k \rangle = \psi(x,k) = \frac{1}{\sqrt{L}}e^{ikx}, \tag{13.1}$$

with corresponding eigenvalues

$$E_0(k) = \frac{\hbar^2 k^2}{2m_0}, \tag{13.2}$$

where $m_0 \sim 9.1 \times 10^{-31}$kg is the free-electron mass. We work in a periodic-boundary condition picture, which requires $\psi(x + L) = \psi(x)$, which requires that the k's are discrete, given by $k_n = (2\pi/L)n$ where n is any integer. We note immediately that a cross-matrix element of the type

$$\langle k_m | k_n \rangle = \int dx \langle k_m | x \rangle \langle x | k_n \rangle = \int dx \psi^\star(x, k_m) \psi(x, k_n) = \frac{1}{L} \int_0^L dx\, e^{i2\pi(n-m)x} = \delta_{n,m} \tag{13.3}$$

is a Kronecker-delta function. This is of course how it should be, since the eigenvectors states $|k_n\rangle$ and $|k_m\rangle$ are mutually orthogonal if (n, m) are different, and the states are normalized to unity. The Hamiltonian matrix is thus diagonal, with the diagonal matrix elements $\langle k | \hat{H}^0 | k \rangle = E_0(k)$ given by the free-electron bandstructure in equation 13.2. The off-diagonal elements $\langle k_m | \hat{H}^0 | k_n \rangle = E_n \langle k_m | k_n \rangle$ are zero because of equation 13.3.

13.3   Periodic perturbation

Now lets add a perturbation to the free electron in the form of a periodic potential. The perturbation potential is

$$W(x) = -2U_G \cos(Gx) = -U_G(e^{iGx} + e^{-iGx}), \tag{13.4}$$

where U_G is the 'strength' in units of energy, and $G = 2\pi/a$, where a is the lattice-constant of the perturbation. The potential is shown in Fig 13.1. The lowest energy of a *classical* particle in this potential landscape is clearly $-2U_G$, at the bottom of a valley. The new Hamiltonian is then $\hat{H} = \hat{H}^0 + W(x) = -\frac{\hbar^2}{2m_0}\frac{\partial^2}{\partial x^2} - 2U_G \cos(Gx)$. In principle this 1D Schrodinger-equation can be solved numerically to a large degree of accuracy *directly* without perturbation theory. But we are going to apply perturbation theory to highlight the insights it affords.

We can find the entire Hamiltonian matrix if we find the matrix elements $\langle k_2 | \hat{H}^0 + W(x) | k_1 \rangle = E_0(k_1)\delta_{k_1,k_2} + \langle k_2 | W(x) | k_1 \rangle$. The first term is the unperturbed diagonal
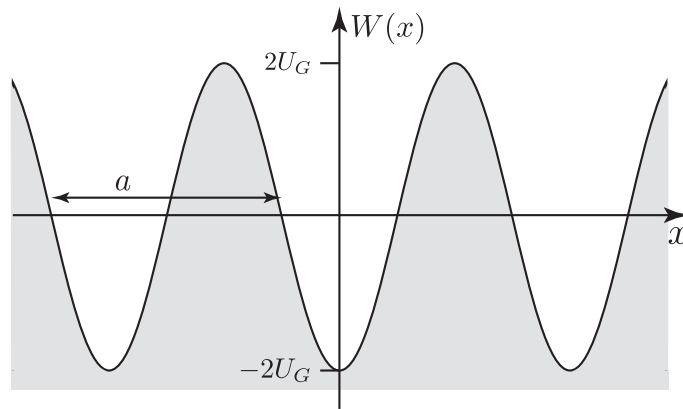
FIGURE 13.1: A periodic potential $W(x) = -2U_G \cos(Gx)$ acts as a perturbation to the free electron.

matrix element, and the second term is due to the perturbation. The perturbation matrix element evaluates to

$$\langle k_2 | W(x) | k_1 \rangle = -\frac{U_G}{L} \int_0^L dx\, e^{i(k_1-k_2)x}(e^{iGx} + e^{-iGx}) = -U_G \delta_{k_1-k_2, \pm G}. \qquad (13.5)$$

The Kronecker-delta implies that the perturbation only couples states $|k_1\rangle$ and $|k_2\rangle$ if their wavevector difference is $k_1 - k_2 = \pm G$, the reciprocal lattice vector of the perturbing potential. Recall from chapter 12 that we can find the perturbed eigenvalues by the matrix method, which works *both* for degenerate and non-degenerate states. But if we were to consider all the $|k\rangle$ states, the matrix would be ∞-dimensional. So we should choose a restricted set for identifying the eigenvalues.

13.4 Degenerate Perturbation Theory

It is clear from equation 13.5 that a state $|k\rangle$ will interact due to the periodic perturbation with only two other states $|k + G\rangle$ and $|k - G\rangle$ directly. This will require us to solve a 3×3 Hamiltonian. But also recall in chapter 12 the result of non-degenerate perturbation theory told us that the changes in eigenvalues for states widely separated in energy goes as $|W_{12}|^2/(E_1 - E_2)$. So the states that interact *most strongly* due to the perturbation must be close (or degenerate) in energy, but their wavevectors should still follow $k_1 - k_2 = \pm G$. Clearly, two such states are $|+G/2\rangle$ and $|-G/2\rangle$. This is illustrated in Fig 13.2. To locate states that have non-zero matrix elements, one has to imagine sliding the double-headed arrow of length G along the k−axis. Two situations are shown, one when the unperturbed states are degenerate, and one when they are not. Also remember the

repulsive nature of the interaction: in Fig 13.2 we expect state $|k_1\rangle$ to be pushed *down*, and state $|k_2\rangle$ to be pushed *up* due to their mutual interaction.

The unperturbed eigenvalue of the two degenerate states is $E_0(G/2) = \hbar^2 G^2/8m_0 = F$. Clearly this is a case for the application of *degenerate* perturbation theory[1]. The problem is rather simple, since the Hamiltonian is a 2×2 matrix:

$$\hat{H}^0 + W = \begin{array}{c} \\ \langle +\frac{G}{2}| \\ \langle -\frac{G}{2}| \end{array} \overset{\displaystyle |+\frac{G}{2}\rangle \quad |-\frac{G}{2}\rangle}{\begin{pmatrix} F & -U_G \\ -U_G & F \end{pmatrix}}, \tag{13.6}$$

where we write out the ket and bra states explicitly to highlight where the matrix elements come from. The eigenvalues of this matrix are obtained by solving the determinant of the matrix: $(F-E)^2 - U_G^2 = 0$, which yields $E_\pm = F \pm U_G$. This implies the degenerate unperturbed states $E_0(+G/2) = E_0(-G/2) = F$ have now been split to two energies E_+ and E_- with the difference $E_+ - E_- = 2U_G$ by the periodic perturbation. This is the opening of a *bandgap* in the allowed energies for the electron, and is highlighted in Fig 13.2.
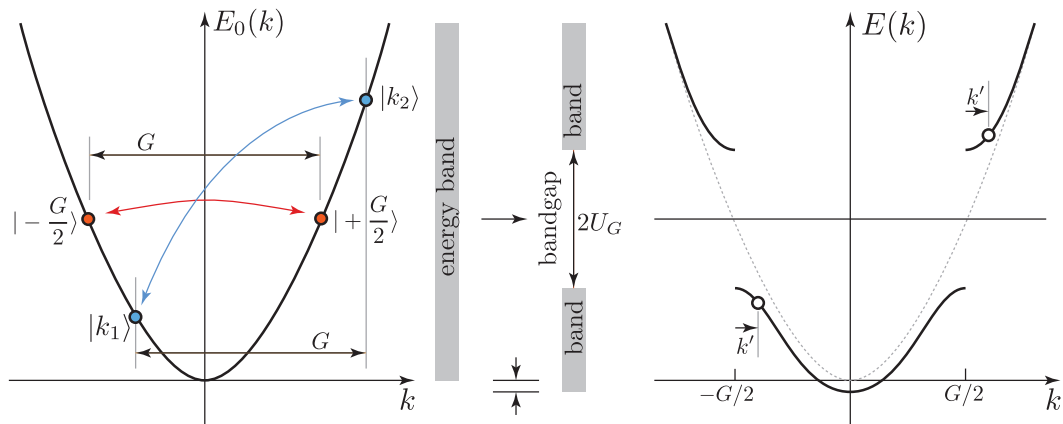

FIGURE 13.2: Bandgap opening in the energy spectrum of a free electron upon perturbation by a periodic potential.

We note here that the general eigenvalues of the 2×2 Hamiltonian matrix

$$\begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \tag{13.7}$$

---

[1]We will see later that the Brillouin-Wigner (BW) non-degenerate perturbation theory also can give the same result.

are

$$E_\pm = \frac{H_{11} + H_{22}}{2} \pm \sqrt{(\frac{H_{11} - H_{22}}{2})^2 + |H_{12}|^2}, \qquad (13.8)$$

the corresponding eigenvectors are

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}_\pm = \begin{bmatrix} \frac{H_{12}}{\sqrt{|H_{12}|^2 + (E_\pm - H_{11})^2}} \\ \frac{E_\pm - H_{11}}{\sqrt{|H_{12}|^2 + (E_\pm - H_{11})^2}} \end{bmatrix} \qquad (13.9)$$

Since we expand the perturbed states as $|\psi\rangle = \sum_n a_n |n\rangle$, for the degenerate states, we obtain the perturbed eigenvectors as

$$|\pm\rangle = a_{1\pm} | + \frac{G}{2}\rangle + a_{2\pm} | - \frac{G}{2}\rangle. \qquad (13.10)$$

For the degenerate states we get $a_{1+} = -1/\sqrt{2}$ and $a_{2+} = +1/\sqrt{2}$, and $a_{1-} = -1/\sqrt{2}$ and $a_{2-} = -1/\sqrt{2}$. The identification of the coefficients helps us convert the perturbed eigenvectors into the eigenfunctions

$$\langle x|+\rangle = \psi_+(x) = (-\frac{1}{\sqrt{2}}) \cdot (\frac{e^{i\frac{G}{2}x}}{\sqrt{L}}) + (+\frac{1}{\sqrt{2}}) \cdot (\frac{e^{-i\frac{G}{2}x}}{\sqrt{L}}) = -i\sqrt{\frac{2}{L}} \sin(\frac{G}{2}x), \qquad (13.11)$$

and

$$\langle x|-\rangle = \psi_-(x) = (-\frac{1}{\sqrt{2}}) \cdot (\frac{e^{i\frac{G}{2}x}}{\sqrt{L}}) + (-\frac{1}{\sqrt{2}}) \cdot (\frac{e^{-i\frac{G}{2}x}}{\sqrt{L}}) = -\sqrt{\frac{2}{L}} \cos(\frac{G}{2}x). \qquad (13.12)$$

This is illustrated in Fig 13.3. Note now the properties of $|\psi_+(x)|^2 = (2/L)\sin^2(Gx/2)$ and $|\psi_-(x)|^2 = (2/L)\cos^2(Gx/2)$. The probability densities for the higher energy states $E_+ = F + U_G$ go as $\sin^2(Gx/2)$, meaning they peak at the highest points of the perturbing potential. The high potential energy is responsible for the high net energy of these states. Similarly, the lower energy states $E_- = F - U_G$ pile up in the valleys, and consequently have lower energies. Note that due to the perturbation, the new eigenfunctions of the degenerate states *no longer* have a uniform probability distribution in space.

But what about states that are not degenerate? Let's look at the states $|k_2\rangle = |\frac{G}{2} + k'\rangle$ and $|k_1\rangle = | - \frac{G}{2} + k'\rangle$, for example those shown in Fig 13.2. By tuning the magnitude
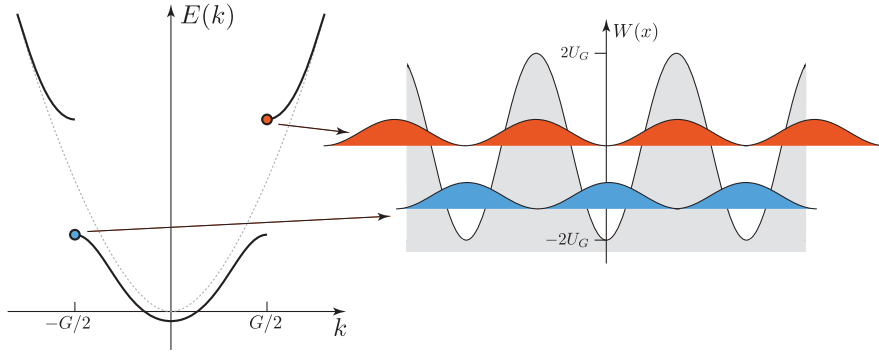
FIGURE 13.3: Probability pileups of band-edge states.

of k', we can move as close to the $\pm G/2$ states as possible. The perturbed Hamiltonian is

$$
\hat{H}^0 + W = \begin{matrix} & |+\frac{G}{2}+k'\rangle & |-\frac{G}{2}+k'\rangle \\ \langle +\frac{G}{2}+k'| \\ \langle -\frac{G}{2}+k'| \end{matrix} \begin{pmatrix} E_0(+\frac{G}{2}+k') & -U_G \\ -U_G & E_0(-\frac{G}{2}+k') \end{pmatrix}, \tag{13.13}
$$

where we write the diagonal unperturbed eigenvalues as

$$
E_0(\pm\frac{G}{2}+k') = \underbrace{\frac{\hbar^2 G^2}{8m_0}}_{F} + \underbrace{\frac{\hbar^2 k'^2}{2m_0}}_{\mathcal{E}(k')} \pm \underbrace{\frac{\hbar^2 G k'}{2m_0}}_{2\sqrt{F\mathcal{E}(k')}} = F + \mathcal{E}(k') \pm 2\sqrt{F\mathcal{E}(k')}. \tag{13.14}
$$

The eigenvalues then are obtained from equation 13.8 as

$$
E_\pm(k') = F + \mathcal{E}(k') \pm \sqrt{4F\mathcal{E}(k') + U_G^2} \approx F + \mathcal{E}(k') \pm U_G(1 + \frac{2F\mathcal{E}(k')}{U_G^2}), \tag{13.15}
$$

where we have expanded the square root term using $(1+x)^n \approx 1 + nx + \dots$ for $x << 1$ assuming $4F\mathcal{E}(k')/U_G^2 << 1$. The energy dispersion then becomes

$$
E_\pm(k') \approx (F \pm U_G) + (1 \pm \frac{2F}{U_G})\frac{\hbar^2 k'^2}{2m_0}, \tag{13.16}
$$

from where we choose the $+$ sign as a 'conduction' band with lowest energy $E_c(0) = F + U_G$, and the $-$ sign as the 'valence' band with highest energy $E_v(0) = F - U_G$. We rewrite the energy dispersions as

$$
\begin{aligned}
E_c(k') &\approx E_c(0) + \frac{\hbar^2 k'^2}{2m_c^\star} \\
E_v(k') &\approx E_v(0) + \frac{\hbar^2 k'^2}{2m_v^\star}
\end{aligned}
\tag{13.17}
$$

where the conduction band *effective* mass is $m_c^\star = \frac{m_0}{1 + \frac{2F}{U_G}}$, and the valence band *effective* mass is $m_v^\star = \frac{m_0}{1 - \frac{2F}{U_G}}$. We note immediately that the *effective* mass of carriers at the band-edges is different from the mass of the free-electron. The conduction band edge effective mass is *lower* than the free electron mass; the electron moves as if it is lighter. If we assume that $U_G << F$, we can neglect the 1 in the denominator, and we get the interesting result that $m_c^\star \sim (U_G/2F)m_0$, that is, the effective mass is *proportional* to the energy bandgap. We will see later in chapter 14 in the $\mathbf{k} \cdot \mathbf{p}$ theory that for most semiconductors, this is an excellent rule of thumb for the conduction band effective mass.

The valence band effective mass under the same approximation is $m_v^\star \sim -(U_G/2F)m_0$, i.e., it is *negative*. This should not bother us at least mathematically, since it is clear that the bandstructure curves *downwards* in Fig 13.2, so its curvature is negative. Physically, it means that the electron in the valence band moves in the opposite direction to an electron in the conduction band in the same $|k\rangle$ state. This is clear from the *group velocity* $v_g = \hbar^{-1} dE(k)/dk$: the slopes of the states are opposite in sign.

Are there $k-$states other than $|\pm G/2\rangle$ at which energy gaps develop due to the perturbation? Let's examine the states $|\pm G\rangle$, with unperturbed energy $4F$. Clearly, $k_2 - k_1 = 2G$, so there is no *direct* interaction between the states. But an *indirect* interaction of the form $|-G\rangle \leftrightarrow |0\rangle \leftrightarrow |+G\rangle$ is possible. This is illustrated in Fig 13.4. The eigenvalues for such an interaction are found by diagonalizing the 3×3 perturbation Hamiltonian

$$
\hat{H}^0 + W = \begin{array}{c} \\ \langle -G| \\ \langle 0| \\ \langle +G| \end{array}
\begin{array}{c}
\begin{array}{ccc} |-G\rangle & 0 & |+G\rangle \end{array} \\
\left(\begin{array}{ccc}
4F & -U_G & 0 \\
-U_G & 0 & -U_G \\
0 & -U_G & 4F
\end{array} \right)
\end{array},
\tag{13.18}
$$

which yields the perturbed eigenvalues $4F, 2F \pm \sqrt{4F^2 + 2U_G^2}$. If the perturbation potential is weak, i. e., $U_G << F$, then we can expand the square root to get the three eigenvalues $4F, 4F + U_G^2/2F, -U_G^2/2F$. We note that there indeed is a splitting of the $|\pm G\rangle$ states, with an energy bandgap $U_G^2/2F$. Similarly, gaps will appear at $\pm mG/2$, due to indirect

interactions $|-mG/2\rangle \leftrightarrow |-(m/2+1)G\rangle \dots \leftrightarrow |+mG/2\rangle$, with a bandgap that scales as U_G^m. For example, the indirect interaction $|-3G/2\rangle \leftrightarrow |-G/2\rangle \leftrightarrow |+G/2\rangle \leftrightarrow |+3G/2\rangle$ is depicted schematically in Fig 13.4.
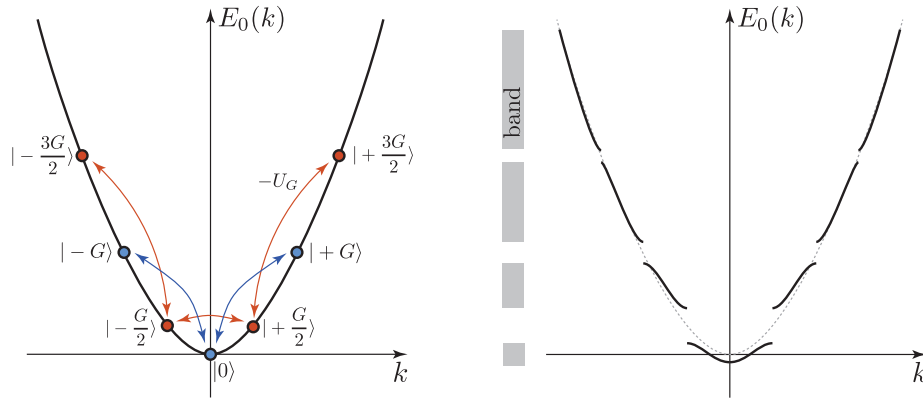

FIGURE 13.4: Indirect coupling via intermediate states. Each coupling has a strength $-U_G$.

We also note that the intermediate state $|k = 0\rangle$ which had a zero unperturbed energy now has been pushed down, and has a negative energy of $-U_G^2/2F$. Thus the ground state energy of the electron is now *negative*, implying it is energetically favorable that the electron be in this state. This idea develops into the concept of a work-function of a solid: it takes energy to kick out an electron from the ground state into the free-electron state, which has a minimum of zero. The work-function is schematically illustrated in Fig 13.2.

13.5   Non-degenerate Perturbation Theory

A number of results that we obtained in the previous section may be obtained using non-degenerate perturbation theory. Recall non-degenerate perturbation theory in chapter 12 Equation 12.23, provided the Brillouin-Wigner (BW) result

$$E \approx E_u + \langle u|W|u\rangle + \sum_{m \neq u} \frac{|\langle m|\hat{W}|u\rangle|^2}{E - E'_m}, \tag{13.19}$$

with the Rayleigh-Schrodinger result obtained by simply replacing $E \to E_u$ on the right side. Let us investigate whether we can apply it to the electron in a periodic potential problem.

If we apply the BW-theory to the states $|u\rangle = |\pm G/2\rangle$, we identify $E_u = F$, $\langle u|\hat{W}|u\rangle = 0$, and $\langle -G/2|\hat{W}|+G/2\rangle = -U_G$, the sum in the RHS of equation 13.19 has just one term, and we get

$$E \approx F + \frac{U_G^2}{E - F} \implies E \approx F \pm U_G, \qquad (13.20)$$

which actually yields the same result as obtained by degenerate perturbation theory in the last section. The two degenerate states are split, with a gap of $2U_G$. This is an advantage of the BW-theory: it works even for degenerate states, though it is typically classified under non-degenerate perturbation theory. Note that we had to solve the same quadratic equation as the 2×2 matrix in the degenerate theory. They are the same thing. The disadvantage of the BW theory is that it requires us to solve for the roots of a polynomial equation.

Clearly the RS-theory

$$E \approx E_u + \langle u|W|u\rangle + \sum_{m \neq u} \frac{|\langle m|\hat{W}|u\rangle|^2}{E_u - E_m'}, \qquad (13.21)$$

cannot be applied to degenerate states, since the the denominator in the sum on the RHS will become zero. But it is well-suited for non-degenerate states. For example, if we ask the question how is state $|0\rangle$ perturbed by its interaction with states $|-G\rangle$ and $|+G\rangle$, we get

$$E \approx 0 + 0 + \frac{U_G^2}{0 - 4F} + \frac{U_G^2}{0 - 4F} = -\frac{U_G^2}{2F}, \qquad (13.22)$$

which is the approximate result we had obtained by diagonalizing the perturbation matrix in equation 13.18. For small perturbations, this result is a good approximation. But if U_G increases, it is easy to see that the minimum energy $-U_G^2/2F$ can become lower than the classically minimum energy allowed in the system, which is $-2U_G$. This should be clear from Fig 13.1. The minimum energy allowed for the electron should be *larger* than $-2U_G$ because of quantum confinement, implying $U_G << 4F$.

Application of the BW theory removes this restriction, since it requires the solution of

$$E \approx 0 + 0 + \frac{U_G^2}{E - 4F} + \frac{U_G^2}{E - 4F} = -\frac{2U_G^2}{E - 4F}, \qquad (13.23)$$

which yields the same result as the non-degenerate matrix method for state $|0\rangle$: $E \approx 2F - \sqrt{4F^2 + 2U_G^2}$. This root is clearly always greater than $-2U_G$, since it asymptotically

approaches $-\sqrt{2}U_G$ when U_G is large. But note that too large a U_G compared to F makes the 'perturbative' treatment not valid in the first place.

13.6 Glimpses of the Bloch Theorem

Consider the free electron state $|k\rangle$ with real-space projection $\psi_k(x) = \langle x|k\rangle = e^{ikx}/\sqrt{L}$. Due to the periodic potential $W(x) = -2U_G \cos(Gx)$, this state couples to the states $|k + G\rangle$ and $|k - G\rangle$. What is the perturbed *wavefunction* in real space?

From Equation 12.22 in chapter 12, we write the perturbed state vector $|k'\rangle$ as

$$|k'\rangle \approx |k\rangle + \frac{\langle k + G|W|k\rangle}{E(k) - E(k+G)}|k+G\rangle + \frac{\langle k - G|W|k\rangle}{E(k) - E(k-G)}|k-G\rangle \qquad (13.24)$$

where we have used the Rayleigh-Schrodinger version. The matrix elements are $-U_G$; projecting the perturbed state on $\langle x|$ we get the perturbed wavefunction to be

$$\psi_{k'}(x) = \langle x|k'\rangle \approx \frac{e^{ikx}}{\sqrt{L}} - \frac{U_G}{E(k) - E(k+G)}\frac{e^{i(k+G)x}}{\sqrt{L}} - \frac{U_G}{E(k) - E(k-G)}\frac{e^{i(k-G)x}}{\sqrt{L}} \quad (13.25)$$

from where we split off e^{ikx} to write the wavefunction as

$$\psi_{k'}(x) \approx e^{ikx} \cdot \underbrace{\left[\frac{1}{\sqrt{L}} - \left(\frac{U_G}{E(k) - E(k+G)}\right)\frac{e^{iGx}}{\sqrt{L}} - \left(\frac{U_G}{E(k) - E(k-G)}\right)\frac{e^{-iGx}}{\sqrt{L}}\right]}_{u_k(x)} \cdot$$

$$(13.26)$$

Note that the wavefunction is of the form $e^{ikx}u_k(x)$, where the function $u_k(x)$ has the property $u_k(x + a) = u_k(x)$, because $e^{\pm iGa} = 1$. This is really the statement of the Bloch theorem: the eigenfunctions for an electron in the presence of a periodic potential can be written in the form $\psi_k(x) = e^{ikx}u_k(x)$, where $u_k(x + a) = u_k(x)$ has the same periodicity as the potential. A more complicated periodic potential such as $W(x) = -2[U_{G_1}\cos(G_1 x) + U_{G_2}\cos(G_2 x) + ...]$ will lead to more couplings, and create more terms in $u_k(x)$, but the Bloch decomposition of the wavefunction in Equation 13.26 will still remain true. We call this a 'glimpse' of the Bloch theorem because of the '$\approx$' sign in Equation 13.26; in the next chapter this sign will be rigorously turned into an equality. The Bloch theorem is a *non-perturbative* result: it does not depend on the strength of

the periodic potential. But of course we just saw it naturally *emerge* as a result from perturbation theory.

13.7 Non-periodic potentials and scattering

We make a few observations of the material covered in this chapter. First, the application of a periodic potential $-2U_G \cos(Gx)$ of reciprocal lattice vector G could only directly couple states that followed $k_2 - k_1 = \pm G$. This caused the appearance of bandgaps due to *direct* interaction at states $|k\rangle = |\pm G/2\rangle$. But due to *indirect* interactions, bandgaps also appeared at $|\pm mG/2\rangle$. If the periodic potential instead was $W(x) = -2[U_{G_1} \cos(G_1 x) + U_{G_2} \cos(G_2 x)]$, we expect direct gaps at more $k-$points, and more direct and indirect coupling of states. The nature of the periodic potential will thus determine the bandstructure.

If instead of a periodic potential, we had a localized potential, say $W(x) = V_0 e^{-x/x_0}$, then we can Fourier-expand the potential to obtain $W(x) = \sum_G U_G \cos(Gx)$, and the expansion coefficients will dictate the strength of the couplings. We immediately note that since a localized potential will require a large number of G's, it will effectively couple a wide range of $k-$states. This is why any *deviation* from periodicity will couple a continuum of $k-$states, a phenomena that is responsible for scattering and localization.

Applications of non-degenerate and degenerate perturbation theory can explain a host of phenomena in semiconductors, and other quantum systems. In this chapter, we applied the techniques to the 'toy-model' of an electron in a 1D periodic potential. In the next chapter, we investigate this technique to develop a rather useful model for the electronic bandstructure of realistic semiconductors.

Debdeep Jena: www.nd.edu/~djena