Part IV

Phonons, Plasmons, Polaritons, and More

Cover Art by Robert Wittig
Front: Chess
Rear: Misty Hills

Preface

The mid-1980's was a vibrant period in my professional life. I had transferred from Electrical Engineering to Chemistry a few years earlier following a six-month sabbatical with Dudley Herschbach. My research group was talented and fearless. We ventured into new areas, worked day and night as if possessed, and good fortune came our way. We made great strides and had a great time. We were happy.

I had taught upper division Physical Chemistry for several years, including thermodynamics, which never warmed my heart. Consequently, I volunteered to take on a graduate course entitled Mathematical Methods of Physical Chemistry, or perhaps it was entitled Mathematical Methods of Chemical Physics. I do not remember, nor is it important. The course had lain dormant since the late 1970's for reasons that were unknown to me at that time, and to a large extent remain unknown to me now. After all, the syllabus had advertised an enticing array of interesting and useful topics. In any event, revival was overdue. Little did I know what was in store. I taught the course more or less without interruption for thirty years.

The audiences in the mid 1980's were entering graduate students in Physical Chemistry, Chemical Physics, and Theoretical Chemistry, plus a sprinkling of students from Physics and Engineering. On occasion a student from Inorganic or Organic Chemistry wandered in, but that was rare and usually the person did not stay long. Topics varied from year to year, drawing from electricity and magnetism, classical mechanics, statistical mechanics, complex variables, linear algebra, Fourier analysis, and the like. My thinking was that such material is essential if one is to deal with quantum mechanics at other than a cursory level, and read journal articles with a critical eye. The course caught on and continued for many years in this mold.

Books and Kettle
Robert Wittig

It was inevitable that the course would evolve in response to the changing times. Some topics were added, some were tossed, and some underwent major revision. Despite the changes, I never wavered in my belief that students entering our program would acquire math skills most readily when the math is applied to physical problems to which they can relate. Besides, I am not a mathematician by any stretch of the imagination, so rigorous proofs were off the table from the outset.

About half a dozen years ago I introduced a series of elementary pedagogical models that describe properties of particles in one-dimensional periodic potentials. These models constitute the ingredients of the first three chapters in Part A. They are designed to elicit features that have three-dimensional counterparts in real systems, for example, band gaps,

Brillouin zones, dispersion relations, metallic versus insulator character, molecular orbitals of periodic systems, density of states, and so on. These models are often encountered in inorganic and materials chemistry, materials science (engineering), and solid-state physics. Many of our entering graduate students had indicated interests along these lines, and it seemed obvious that this trend would continue and most likely accelerate.

Chapter 4 of Part A is an introduction to the usual solid-state physics approach centered on the use of Bloch functions for describing electrons in periodic lattices. Chapter 5 is relatively short: a glimpse into the realm of molecular crystals through the enlistment of a strategy espoused by Roald Hoffmann that is akin to that of the Hückel model. Part A was typically covered in a month or so. The material is straightforward but requires a lot of work: a jumpstart, so to speak, for students entering our program. There are exercises, classroom participation, a couple of quizzes, and even a couple of student presentations. One of the main goals of Part A is to set the stage for Part B, which is the larger and more interesting part.

Part B comprises the main topics covered in the course. The goal is to endow the students with a reasonable understanding of properties and phenomena that manifest in solids, both large and small. The journey begins with lattice vibrations. Quantization of a discrete mass field yields phonons. The character of these massless bosons emerges straightforwardly via *k*-space quantization using commutators. The Hamiltonian, when represented in *k*-space, acquires a transparent Fock space representation. This stuff is really neat. It has far-reaching implications that are encountered later in Part B, as well as in the very different context of relativistic quantum field theory in Part V of the Notes Project.

Ink Well and Glasses
Robert Wittig

Chapter 2 covers heat capacity and thermal conductivity, which are less boring than one might imagine. From there we head toward plasmonics, polaritons, and applications. Chapters 3 and 4 introduce many aspects of insulators, metals, and semiconductors. We see how screening in metals enables a free-particle picture to be applicable in the face of strong electron correlation. Plasmas in metals and semiconductors are discussed: propagation, quanta of the plasma field that are referred to as plasmons, wave vector conservation, surface plasmons, charge fluctuations, applications, and more. Chapter 5 continues in this vein. It introduces the principles that underlie nanoscale plasmonic devices, emphasizing the benefits that accrue when working with particles whose dimensions are significantly smaller than the radiation wavelength. This material is timely, and it relates to a large amount of research in our department and other departments at USC.

The fascinating topic of polaritons is discussed in Chapters 6 and 7. Systems that carry electric charge, by definition, interact with electromagnetic fields. Consequently, the elementary excitations of such systems are combinations of the coherent motions of the charged constituents and the electromagnetic fields thereby engendered. A related and

conceptually appealing example is a polaron. This is an electron moving through a crystal, say, by hopping from site to site. Of course, the electron polarizes its immediate surroundings, so its transport is coupled to that of its polarization cloud. The same general idea applies to phonons, plasmons, and other such field quanta despite the fact that they are bosons, whereas the polaron is a fermion. Namely, they are all inextricably coupled to electromagnetic fields.

Chapter 7 describes plasmon polaritons that exist at a dielectric-metal boundary. The dispersion relation shows two branches. The more interesting one describes an evanescent wave that propagates along the surface, with electromagnetic fields that decay exponentially into the dielectric and metal regions on either side of the interface. Assuming minimal loss, there is a well-defined forbidden gap where propagation cannot take place. This aspect is akin to the phonon polariton.

A fairly dramatic switch of emphasis takes place in Chapters 8 and 9. We turn away from the fields and their quanta that arise with coupled particle ensembles: mass points and springs, plasmas in metals and semiconductors, nanoscale plasmonics, and phonon and plasmon polaritons. Additions and changes include a trivial extension of the dual-use field quantization model introduced in Chapter 1 from one to three dimensions, letting the volume go to infinity, and appending explicit time behavior to the field operator. These moves are not without purpose. They are prerequisite to a field theory model that comes into its own in Part V. The end product of Chapter 8 is a field operator that creates and annihilates quanta of bosonic fields.

Something very neat then happens. Chapter 9 starts with the model that entered way back at the beginning of Chapter 1. The mathematical form of its dispersion relation is retained, the constants are redefined, and classical variables are replaced with counterpart operators. The resulting d'Alembertian operator acts on a function that is interpreted as the scalar field of a relativistically covariant theory. These maneuvers yield the Klein-Gordon equation of relativistic quantum field theory. As mentioned above, this stuff comes into its own in Chapter 4 of Part V.

Hot Stuff
Robert Wittig

These notes are a slightly better organized version of the ones that were distributed in class during the past half dozen years. They are intended to complement lectures and textbooks, not replace them. There is obviously far too much material to be covered in one semester. This was never even attempted. Emphasis varied from year to year, depending on student interests, backgrounds, and preparedness. We generally covered about half of the material contained herein in any given semester, though nearly all of it saw action during the past half dozen years.

It needs to be emphasized that these notes do not constitute a textbook despite their formal veneer. The style is different and the delivery is relatively inhomogeneous: sometimes mathematical, sometimes qualitative, parts unfinished or in need of help, repetition, and things are occasionally out of order. Topics were discussed in class that are not con-

tained herein, but would have been if this were a textbook. Perhaps the best thing about the notes is that the person who wrote them is the person who gave the lectures. This synchrony differs from teaching a class using assigned texts, however well written. It catalyzed classroom discussion, as evidenced by the large volume of questions during class – every teacher's holy grail. In my opinion, these notes are infinitely superior to distributing copies of slides prior to lectures, a practice that I find utterly useless.

Students who take the class seriously and are diligent insofar as the assigned work are rewarded. It opens their eyes to research areas they are destined to encounter. For example, the material shows up frequently in our Physical and Theoretical Chemistry Seminar Series. The students also find that the course helps (and sometimes influences) them when the time comes to choose a research direction and advisor. They might be unaware of this influence, but it is there nonetheless.

Curt Wittig
August 2015

Contents: Part A

Contents: Part B

Part A

Chapter 1.

Particle-on-a-Ring

White Houses
Robert Wittig

"In answer to the question of why it happened, I offer the modest proposal that our universe is simply one of those things that happen from time to time."

Edward P. Tryon

Contents

Sailboat
Robert Wittig

The Road goes ever on and on
Down from the door where it began,
Now far ahead the Road has gone,
And I must follow if I can,
Pursuing it with eager feet,
Until it joins some larger way
Where many path and errands meet.
And whither then? I cannot say.

The Lord of the Rings
J. R. R. Tolkien

Preliminary Comments

The subject of particles, say electrons, in periodic potentials is examined in this and the next three chapters. The goal is to gain understanding through toy models. Previous exposure is not required, though it is assumed that the reader is comfortable with quantum mechanics. We shall start with one-dimensional (1D) models having modest numbers, N, of evenly spaced sites, and build from there, eventually reaching the $N \to \infty$ limit. It is easy to vary shapes and strengths of potentials to access different regimes. A more formal approach is introduced in Chapter 4. Connections between small-N and large-N regimes are noted throughout Part A. It is uncanny that so many results obtained via solid-state physics (energy bands, band gaps, Bloch waves, Brillouin zones, dispersion curves, nearly free electron limit, tight binding limit) are intuitive extensions of small-N models to a large numbers of sites. This paves the way to the crystalline phase.

Consider a 1D potential $V(x)$ defined in the region $0 \le x \le a$. Its infinite 1D lattice counterpart is obtained by repeating this segment *ad infinitum* through the condition $V(x + na) = V(x)$, where n is an integer. This yields a lattice of evenly spaced identical intervals. Electrons in this lattice can be nearly free, bound strongly to the sites, or lie between these extremes, depending on how many of them are present, and where their energies lie with respect to the barriers that separate the sites. Though we live in a 3D world, a 1D lattice using simple potentials will be used to illustrate phenomena. Think of this as practice *en route* to the more interesting but more complicated 3D systems. The lattice periodicity itself results in interesting phenomena. Indeed, many results are independent of details of the potential.

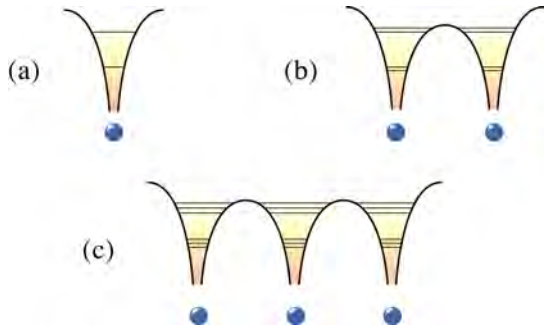


Figure 1. (a) An isolated atom has well-defined levels and ionization potential. (b) Levels are near-degenerate at large R, but as R decreases, degeneracy is lifted. (c) With 3 atoms, triple degeneracy is lifted, and so on for more atoms. The number of closely spaced levels increases linearly with the number of atoms, eventually forming bands.

Referring to Fig. 1(a), the potential experienced by an electron attracted to the atomic ion core supports bound orbitals. The atom's occupied orbitals are indicated with horizontal lines. The ion core is represented by a blue dot. A simple 3D example is a ground state alkali atom. In this case, the highest occupied orbital is that of the outermost s-electron, and the energy difference between the energy of the ground state and the large-r horizontal asymptote is the ionization potential.

In Fig. 1(b), a second atom is brought in. At large interatomic distance, R, the orbitals of the atom pair are, for all practical purposes, doubly degenerate (electron and nuclear coordinates are labeled with lower and upper case letters, respectively). When R is made smaller, degeneracy is lifted because the electrons move in the potential of the incipient

diatom. For three atoms, triple degeneracy at large R is lifted as the atoms are brought together, as indicated in (c). As yet more atoms are brought in, the number of closely spaced orbitals that share a common atomic parentage increases. For N evenly spaced atoms, each atomic orbital gives rise to N molecular orbitals. Spin doubles this: $2N$ spin-orbitals in each group that arises from a common atomic orbital.

If the number of atoms is sufficiently large, the discrete orbitals merge into bands (Fig. 2). Each has its parentage in one of the atomic levels. In general, the lower the energy of the band the narrower its width, because low energy atomic orbitals are perturbed less by neighboring atoms than are their higher energy counterparts. In 3D, bands sometimes overlap because the fields created by the ionic lattice are not the same in different directions. In 1D, however, bands cannot overlap. This is akin to the non-crossing rule in diatomic molecules.

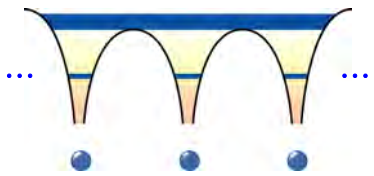


Figure 2. For N atoms, where N is large, the number of discrete orbitals that have energies close to one another is large. These orbitals merge into bands.

The number of electrons in the various bands follows the usual rule for placing fermions in orbitals. Thus, the low energy bands are fully occupied. These electrons are influenced by their respective ion cores (as well as by the other electrons that are strongly bound to the same ion core) much more than by adjacent atoms. Therefore, they resemble their atomic counterparts, and their bands are narrow. The highest energy electrons live in the uppermost occupied band. If each atom contributes an electron to this band, as with alkalis, the band has N electrons, and therefore it is half full.

Roughly speaking, this enables two extremes to be understood. When a band is half full, electrons move about freely because many unoccupied orbitals are accessible at essentially the same energy. Thus, we have an electrical conductor. Alternatively, if each atom contributes two electrons to the uppermost band, the band is full. In this case we have an electrical insulator. Reality is more complicated, but this picture provides a conceptual basis for understanding the difference between conductors and insulators.

The Schrödinger equation will be solved for periodic potentials having small numbers of repeating units. The number of units will then be increased to anticipate the limit of an infinite lattice. This is done first with a particle-on-a-ring model and a computer program that diagonalizes a Hamiltonian matrix. Next, potentials having arbitrary shapes are analyzed by compounding the multiplication of transfer matrices to evolve the system between sites. Connections between small and large numbers of repeating segments are also examined using Hückel theory.

We then use a formal approach based on discrete translational invariance and Fourier analysis. The Schrödinger equation is examined for limiting cases of nearly free and tightly bound electrons. Finally, we return to the molecular perspective, but with a broader view than that of Hückel theory. Bloch's theorem yields the group theoretical results of small symmetric systems such as Na_3, cyclobutadiene, and benzene. Orbital instability known as Jahn-Teller (Peierls) distortion is identified. Extensions are noted.

Part A is repetitious. The intent is to make the case from complementary perspectives. You may skim parts but be careful, as much of the material reappears in Part B.

1.1. Introduction

The goal of Part A is qualitative understanding of systems in which particles experience periodic potentials. It is not a bad idea to think of the particles as electrons, albeit with the understanding that we are going to forego electronic structure issues. When the number of repeating units is large, we have in effect a 1D crystal. It is important to see what happens in the nanoscale regime, where traits of crystalline and large-molecule behavior can coexist, depending on the number of binding sites and the shape and strength of the potential. For example, it is instructive to examine length scales in the range 1-100 nm. Toy models will be used, as their ingredients are transparent and they can be applied to regimes of interest. The strategy that underlies Part A is to use easy models as a first step toward unraveling more complex systems.

In the present chapter, phenomena that manifest in periodic lattices will be examined. Progression from few to many repeating sites will be examined using the particle-on-a-ring model. This is one of the simplest 1D models in all of quantum mechanics. It will be seen that rings that contain 3, 6, 12, and 24 sites illustrate interesting effects as well as emerging traits of infinite 1D lattices. Figure 3 gives an example of a periodic potential. Despite its pathological shape, it is useful as a means of introducing concepts.

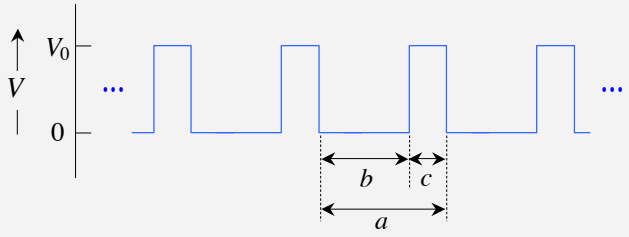


Figure 3. One-dimensional rectangular potential: The length of a well is b, the barrier length is c, and the length of the unit cell is $a = b + c$. A particle-on-a-ring model will be discussed, so a, b, and c will be expressed in terms of distances or angles, depending on context.

Identical Sites

The particle-on-a-ring (POR) model is of great pedagogical value. On rare occasion a real system can be modeled in first approximation using a POR model. However, the model's value lies with the introduction of concepts. For example, in elementary quantum mechanics it is used to introduce angular momentum, degeneracy, and periodic boundary conditions. In the present chapter, single-particle eigenvalues and eigenfunctions will be obtained for a periodic potential $V(\phi)$ on a ring. The potential can be weak or strong and the number of sites can be few or many. The math is the same: matrix diagonalization. Expansion of $V(\phi)$ in a Fourier series and use of an $e^{im\phi}$ basis enable the Hamiltonian matrix to be written straightaway. By adjusting parameter values, the system is endowed with traits associated with core levels, valence bands, conduction bands, and so on, admittedly in 1D. Results obtained in subsequent chapters will be anticipated.

A complementary strategy is used in Chapter 2: Transfer Matrices. The facts that $V(\phi)$ is periodic and wave functions satisfy the cyclic boundary condition $\psi(0) = \psi(2\pi)$ enables us to relate wave amplitudes at different sites to one another using nothing more than phase factors. For example, wave amplitudes at adjacent sites can be taken as identical in every respect except that they differ by $e^{i\delta}$. When there are N sites, the cyclic boundary condition $\psi(0) = \psi(2\pi)$ requires that the round trip phase, $N\delta$, is equal to 2π times an integer. The eigenvalue problem is reduced to a calculation carried out using a single site. Regardless of the shape of the potential for one interval, as long as it can be represented by a piecewise constant approximation, the problem can be solved straightforwardly.

The potential in Fig. 3 is easily expanded in a Fourier series. The fact that all periodic potentials can be expanded in Fourier series leads to a simple algorithm. Matrix elements of terms in the expansion are easily evaluated and their roles are transparent. The shape of a potential can be manipulated without affecting the methodology simply by choosing Fourier coefficients. For small V_0, the functions $e^{im\phi}$ are an intuitive basis because the potential perturbs an otherwise free particle on the ring. However, even with small V_0, the potential efficiently couples degenerate basis functions whose periodicities act in concert with the periodicity of the potential. When these conditions are satisfied, the eigenfunctions are standing waves. This resonance and its near-resonance counterparts enjoy a privileged status in the relationship between small numbers of binding sites and the infinite lattice.

Except for the above resonance, a weak potential has no significant affect on the basis functions, and traveling waves $e^{im\phi}$ serve as acceptable descriptors (zero-order states). In the infinite lattice, this regime is referred to as the nearly-free-electron limit. As mentioned above, it will be shown in the present chapter that upon increasing the number of sites from 6 to 12 to 24, the system progresses toward the infinite lattice limit.

On the other hand, 3 sites gives no indication of infinite lattice behavior because the cyclic boundary condition (ring circumference = integral number of wavelengths) is incompatible with standing waves at the lowest resonance. This requires spacing between adjacent sites of a half wavelength. This is unimportant when the number of sites is large, but it is important when the number of sites is small. For example, you will see in Chapter 3 that small numbers of sites display interesting and important behavior in the Hückel model of conjugated hydrocarbon networks. If you have had a course in organic chemistry, you have already encountered the Hückel model. It will be used to examine conjugated hydrocarbons: open and closed chains, and aromatics ranging from benzene to graphene. Jahn-Teller distortion will be introduced at a qualitative level.

The picture changes qualitatively for large V_0. The orbitals that constitute the lowest band now have the electrons bound strongly to the wells. These orbitals are nearly degenerate because there is little communication between the wells. The energy width of the band is small because it depends on the extent to which there is communication – the smaller the degree of communication, the narrower the band. You might find it useful to think about core electrons of atoms in a monatomic crystal. These electrons care little about their core counterparts on adjacent atoms. The orbitals in the next highest band resemble the first excited state of a particle in a well, and so on for progressively higher

bands until the electrons are able to move rather freely between the wells. It will prove interesting to vary parameters to explore this progression.

1.2. Weak V(ϕ)

For small $V(\phi)$ values, a Hamiltonian matrix of modest dimension (represented in the $e^{im\phi}$ basis) suffices. Matrix diagonalization enables us to see how the eigenvalues and eigenfunctions depend on the model's parameters. Diagonalization is a brute-force approach: inelegant, always works, and often gives results whose interpretation is challenging. However, it is easy. Later it will be shown that the small $V(\phi)$ regime can be modeled accurately using nothing more than 2×2 matrices.

To begin, recall the eigenfunctions and eigenenergies for a ring with $V(\phi) = 0$:

$$\psi_m = e^{im\phi}(2\pi)^{-1/2} \tag{1.1}$$

$$E_m = Bm^2, \tag{1.2}$$

where m is the quantum number. The rotational constant B is equal to $\hbar^2 / 2m_e r_0^2$, where r_0 is the radius of the ring and m_e is taken to be the mass of an electron. The functions ψ_m shall be used as basis functions for examination of the weak $V(\phi)$ regime. They can serve as a basis even in the strong $V(\phi)$ regime as long as a sufficient number of basis functions are used. The matrix elements of $V(\phi)$ are given by

$$V_{m'm} = \langle m' | V(\phi) | m \rangle = \frac{1}{2\pi} \int_0^{2\pi} d\phi \; e^{-im'\phi} V(\phi) e^{im\phi} \tag{1.3}$$

The mathematical procedure will be illustrated using a 6-fold potential. Referring to Fig. 3, we shall use $b = c$, with the horizontal axis defined such that $V(\phi)$ is odd about $\phi = 0$. The horizontal axis could also be defined such that it is even about $\phi = 0$. The expansion of $V(\phi)$ is a textbook example of a Fourier series. If you wish, Fourier analysis will be discussed during one of our Friday sessions. However, I expect you to achieve a reasonable level of familiarity and comfort with Fourier analysis on your own.

The expansion of $V(\phi)$ over the range $0 \le \phi \le 2\pi$ is given by

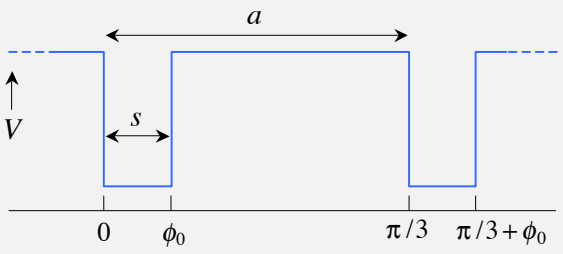
$$V(\phi) = \underbrace{\tfrac{1}{2} V_0}_{\bar{V}} + \sum_{n=1,3,\ldots} B_n \sin 6n\phi, \tag{1.4}$$

where $\bar{V}$ denotes the average value of the potential.

The sum in eqn (1.4) excludes even n because of the symmetry of the potential. Referring to Fig. 3, for even n, integrating $B_n \sin 6n\phi$ over each flat region returns zero because an integral number of wavelengths are contained in each flat region of angular interval $\pi/6$. Thus, the integral that gives B_n [*vide infra*, eqn (1.5)] vanishes.

Were a situation chosen in which $b \neq c$ (see Fig. 3), even and odd n would have nonzero coefficients. For example, consider the $V(\phi)$ shape indicated in Fig. 4. In this case, it is not possible to express the potential as odd with respect to some point. The potential can be represented as an even function, but not as an odd function. Therefore, let us consider an expansion in terms of $A_n \cos 6n\phi$ The ratio a/s gives the smallest value of n that is absent in the expansion. For example, $a/s = 2$ eliminates 2, 4, 6, and so on, whereas $a/s = 3$ eliminates 3, 6, 9, and so on. For large values of $2\pi/\phi_0$, there are, in effect, no missing terms in the expansion of $V(\phi)$. Of course, when a/s is not an integer, there are no missing terms.

Figure 4. The length of the rectangular well is s, and its angular spread is $s/r_0 = \phi_0$. The length of a repeating segment is a. For 6-fold $V(\phi)$ the angular spread of a repeating element is $\pi/3$. Integer values of a/s result in missing n values in the expansion of $V(\phi)$.



Referring to eqn (1.4), for the $\frac{1}{2}V_0$ term, only its diagonal matrix elements are nonzero. Each has the same value: the average value of the potential, $\bar{V}$, which is an unimportant offset that can be suppressed if we so wish. The expansion coefficients B_n are

$$B_n = \frac{1}{\pi}\int_0^{2\pi} d\phi\, V(\phi)\sin 6n\phi\,. \tag{1.5}$$

Evaluating the integral yields

$$B_n = -\frac{2V_0}{n\pi}, \tag{1.6}$$

where $n = 1, 3, 5$, and so on. Keep in mind that $B_n = 0$ for even n because we have set b equal to c in Fig. 3, and labeled the horizontal axis such that only odd terms are nonzero. The minus sign on the right hand side of eqn (1.6) is not important. For example, shifting the origin by $\pi/6$ would give $B_n = +2V_0/n\pi$, despite nothing having changed physically. Putting the above expression for B_n into eqn (1.4) yields

$$V(\phi) - \bar{V} = -\sum_{n=1,3\ldots} \frac{2V_0 \sin 6n\phi}{n\pi}. \tag{1.7}$$

Using complex exponentials [which facilitate solution of eqn (1.3)], eqn (1.7) becomes

$$V(\phi) - \bar{V} = \sum_{n=1,3,..} \frac{iV_0}{n\pi}\left(e^{i6n\phi} - e^{-i6n\phi}\right). \tag{1.8}$$

Inserting the right hand side of eqn (1.8) into eqn (1.3) yields the matrix elements $V_{m'm}$, rather few of which are nonzero. Only basis functions for which $|m'-m|$ is an odd integer multiple of 6 have nonzero $V_{m'm}$ matrix elements.

Now consider the first nonzero term (leaving aside $\bar{V}$) in the expansion of $V(\phi)$ given by eqn (1.4), namely, $B_1 \sin 6\phi$. For this term to have nonzero matrix elements requires that $|m'-m| = 6$. In this case, ψ_m can couple to $\psi_{m'}$ for m and m' values that satisfy $m'-m = \pm 6$. Matrix elements of the next nonzero term in the series ($B_3 \sin 18\phi$) vanish unless $|m'-m| = 18$, and so on for higher terms in the expansion of $V(\phi)$. The n^{th} term in the expansion couples states for which $6n = |m'-m|$.

Basis functions with $m' = -m$ (or close to it) are particularly important. Mixing is 100% when the condition $6n = |m'-m|$ is met because ψ_m and $\psi_{m'=-m}$ have the same unperturbed ($V_0 = 0$) energy. For example, for $n = 1$, the matrix element $V_{m'm}$ that couples $m = 3$ and $m' = -3$ is obtained by putting $-i(V_0/\pi)e^{-i6\phi}$ from eqn (1.8) into eqn (1.3). This yields $V_{-3,3} = -iV_0/\pi$. Thus, the magnitude of the energy splitting is $|2V_0/\pi|$ in the weak $V(\phi)$ limit.

The requirements of degeneracy and $|m'-m|$ being an odd integer multiple of 6 are satisfied for $|m| = 3$, 9, and so on. The strongest coupling takes place with pairs such as (3,–3), (9, –9), and so on. The degenerate pair (6, –6) is weakly coupled because coupling only enters in second order. For example, $V(\phi)$ acting on $m = 6$ creates amplitude at $m = 0$ and 12, and $V(\phi)$ then acting on $m = 0$ creates amplitude at $m = -6$.

Wave Vector

The above equations have been expressed in terms of the angular variable ϕ. If we wish to express the phase in terms of distance x along the ring, the phase $m\phi$ in $e^{im\phi}$ becomes

$$m\phi = mx/r_0, \tag{1.9}$$

where r_0 is the ring's radius. For a potential with 6 identical segments, each of length a, the circumference is $L = 6a = 2\pi r_0$. Using $m\phi = kx$, the wave vector k is given by

$$k = \frac{m\phi}{x} = \frac{m}{r_0} = m\frac{2\pi}{L}. \tag{1.10}$$

For m values that satisfy the exact resonance requirements: $|m'-m| = 6n$ and $m' = -m$ (m/m' values of $\pm 3, \pm 6, \pm 9\ldots$ in the present example), eqn (1.10) gives the corresponding k values: $\pm\pi/a, \pm 2\pi/a, \pm 3\pi/a\ldots$.

The special values of k that correspond to the resonances ($\pm\pi/a, \pm 2\pi/a, \pm 3\pi/a \ldots$) are independent of the number of segments on the ring. To see why this is true, consider eqn (1.10), but now put 24 segments on the ring instead of 6. In other words, use $L = 24a$ instead of $6a$. The requirements: $|m' - m| = 24$ and $m + m' = 0$, result in the smallest value of $|m|$ that satisfies these conditions being 12. However, this corresponds to $k = m(2\pi/L) = \pm\pi/a$. All that has changed is the value of $|m|$ that is needed to satisfy the resonance conditions. This occurs when the spacing between adjacent sites is equal to an integer number of half wavelengths. Thus, using $k = 2\pi/\lambda$ with $\lambda = 2a$ yields: $k = \pm\pi/a$, $\pm 2\pi/a, \pm 3\pi/a \ldots$. These values of k define what is called Brillouin zones.

Figure 5 illustrates the phase progression and wave vector indicated in eqns (1.9) and (1.10) upon going from $m = 1$ to $m = 3$. Note that $m = 1$ corresponds to one wavelength in the angular range $0 - 2\pi$ (equivalently, in the distance range $0 - L$). The other resonances that occur at larger $|k|$ values will be present in proportion to the magnitudes of the corresponding Fourier components of the potential.

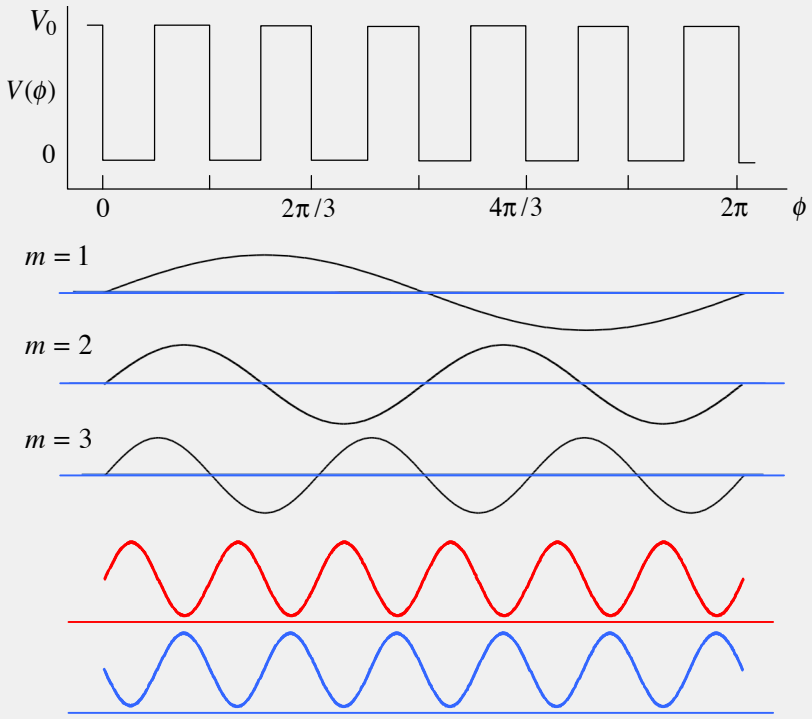


Figure 5. The potential $V(\phi)$ has sixfold symmetry. The plots labeled $m = 1, 2$, and 3 are the imaginary parts of the functions $e^{im\phi}$. They illustrate phase progression. Taking combinations of $e^{i3\phi}$ and $e^{-i3\phi}$ yields standing waves. For the given potential the appropriate standing wave pair is $\sin 3(\phi + \pi/12)$ and $\cos 3(\phi + \pi/12)$. The squares of these yield probability densities that peak over the low and high parts of the potential, respectively. These densities are shown as the red and blue curves. Had the horizontal axis been labeled such that $V(\phi)$ is an even function of ϕ (shift the labels on the ϕ axis to the right by $\pi/12$), the expressions for the standing waves would not have required the $\pi/12$ phase shift. The red and blue curves would then be $\cos^2 3\phi$ and $\sin^2 3\phi$, respectively.

1.3. Diagonalization

Eigenvalues are obtained for n-fold periodic potentials, beginning with rectangular ones having $a/s = 2$. For the six-fold potential in Fig. 5, matrix elements $V_{m'm}$ are evaluated using eqns (1.3) and (1.8) and added to the diagonal elements given by eqn (1.6). The resulting matrix is

m ⟵ ⟶

		−6	−5	−4	−3	−2	−1	0	1	2	3	4	5	6	
		↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	
−6	←	36B	0	0	0	0	0	M	0	0	0	0	0	0	→
−5	←	0	25B	0	0	0	0	0	M	0	0	0	0	0	→
−4	←	0	0	16B	0	0	0	0	0	M	0	0	0	0	→
−3	←	0	0	0	9B	0	0	0	0	0	M	0	0	0	→
−2	←	0	0	0	0	4B	0	0	0	0	0	M	0	0	→
−1	←	0	0	0	0	0	B	0	0	0	0	0	M	0	→
0	←	M^*	0	0	0	0	0	0	0	0	0	0	0	M	→
1	←	0	M^*	0	0	0	0	0	B	0	0	0	0	0	→
2	←	0	0	M^*	0	0	0	0	0	4B	0	0	0	0	→
3	←	0	0	0	M^*	0	0	0	0	0	9B	0	0	0	→
4	←	0	0	0	0	M^*	0	0	0	0	0	16B	0	0	→
5	←	0	0	0	0	0	M^*	0	0	0	0	0	25B	0	→
6	←	0	0	0	0	0	0	M^*	0	0	0	0	0	36B	→
		↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	

(1.11)

where B is the rotational constant and $M = -iV_0/\pi$.

Off-diagonal elements from terms in the expansion of $V(\phi)$ that satisfy $|m'-m| = 6n$ ($n = 3, 5, 7\ldots$) lie outside the range shown in eqn (1.11). They play similar roles to M and M^*. Namely, for each allowed $|m'-m|$ value, strips of matrix elements analogous to the ones in eqn (1.11) are present in the matrix. The approach works with potentials having other shapes, for example in Fig. 6, where the ratio a/s is no longer meaningful.

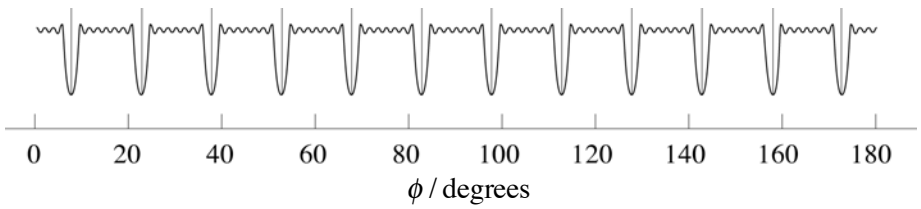


Figure 6. This potential was obtained using 12 Fourier coefficients.

The potential in Fig. 6 has nonzero off-diagonal matrix elements at $|m'-m| = 6$, 12, 18, etc. Yet, except for the missing resonances in the case of the potential in Fig. 5, no qualitative difference arises between potentials such as those shown in Figs. 5 and 6 in the weak $V(\phi)$ limit. Realistic potentials will not have missing resonances. Some Fourier expansion coefficients might be small, but they will not be zero without a rigorous symmetry. Nonetheless, the missing resonances that arise with the potential in Fig. 5 illustrate the role of the Fourier coefficients. It should be obvious that $V(\phi)$ of any shape can be handled: Evaluate its Fourier components and put strips into the Hamiltonian matrix.

In eqn (1.11), the degenerate pair $m = \pm 3$ is indicated with red ovals at their diagonal matrix elements, and blue ovals at their off-diagonal matrix elements, highlighting exact resonance. The arrows at the edges of the matrix indicate that the matrix continues to whatever size is needed.

n-Fold Potentials

Figure 7 shows results for 6, 12, and 24 evenly spaced sites on the ring. The matrices for 12 and 24 sites are the same as eqn (1.11) except that the M and M^* strips lie further from the diagonal. Only positive m values are used for the horizontal axis, as there is mirror symmetry about the vertical axis. In these calculations, V_0 is sufficiently small that a significant percentage of the eigenstates resemble the POR basis. This enables meaningful plots to be made in which the horizontal axis is labeled with the basis index m. Namely, the eigenstates resemble either the $e^{im\phi}$ basis, or a combination of just two functions $e^{im\phi}$ and $e^{im'\phi}$, where m and m' satisfy $|m'-m| = 6n$. Thus, m' values that satisfy $|m'-m| = 6n$ can be added, as indicated in Fig. 7(c). Keep in mind that Fig. 7 is germane to the low-V_0 limit. For larger V_0 there is significant site localization, and eigenstates comprise many $e^{im\phi}$ basis functions.

It is easy to identify the resonance regions centered at $m = 3$, 6, and 12, in (a), (b), and (c), respectively (black dots). At and near exact resonance (black dots), m is not an approximately good quantum number. Pairs of states for which $|m'-m|$ is commensurate with a term in the Fourier expansion of $V(\phi)$ are strongly coupled when $|E_m - E_{m'}|$ is comparable to or smaller than the magnitude of the coupling matrix element. This yields eigenstates that are comprised almost entirely of ψ_m and $\psi_{m'}$. Other basis functions are present in the eigenfunctions, but only to a minute degree. Professor George Kumi (Rutgers University) prepared Figs. 7–12 when he was a graduate student here.

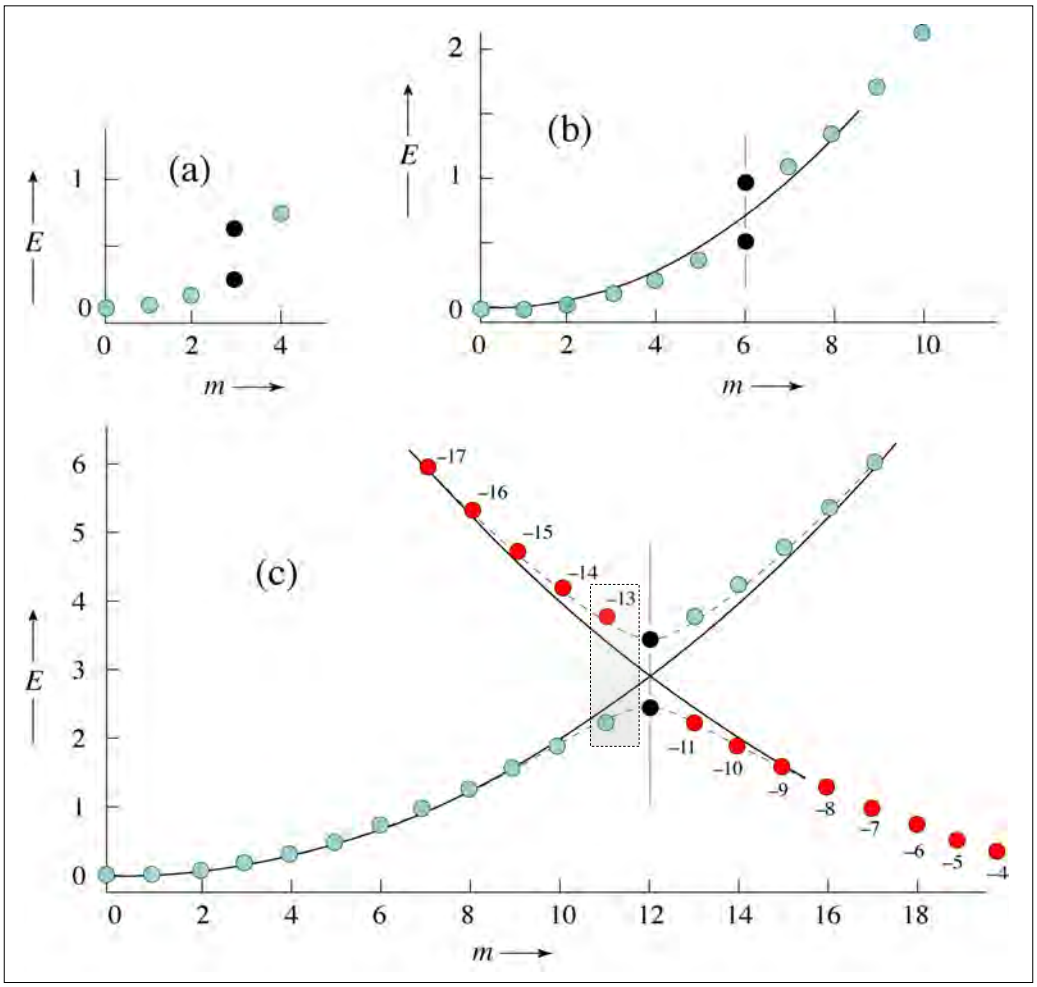


Figure 7. Eigenvalues for 6, 12, and 24-fold $V(\phi)$: Solid curves are parabolas. In (c), the eigenstates are color-coded to highlight POR nature. At exact resonance (black points): $|m'-m| = N$ and $m = -m'$. Near the exact resonances, m is not an approximately good quantum number, because the potential efficiently couples pairs of basis functions for which $|m'-m| = N$. When $V(\phi)$ is weak, eigenstates near the resonances consist of essentially two basis functions. For example, the shaded box indicates the combination of ψ_{11} and ψ_{-13}. The numbers above or below the red points indicate the m' values that go along with the m values on the horizontal axis. Dashed curves indicate "avoided crossing." The red points have been shifted from the left by $\Delta m = 24$.

Pairwise Interaction

Consider the pairs m/m' that satisfy $|m'-m| = 24$. In the small $V(\phi)$ regime, 2×2 matrices suffice to describe the interactions. For example, though ψ_{11} has nonzero matrix elements with both ψ_{-13} and ψ_{35}, the latter is unimportant because $|E_{11}-E_{35}|$ is very much larger than $|E_{11}-E_{-13}|$. Referring to Fig. 7(c), levels with $m \sim 12$ can there-

fore be treated by considering resonating pairs in which $m' \sim -12$, with the understanding that $|m' - m| = 24$. The Hamiltonian matrix for each such pair

$$\begin{pmatrix} Bm^2 & -i(V_0/\pi) \\ i(V_0/\pi) & Bm'^2 \end{pmatrix} \tag{1.12}$$

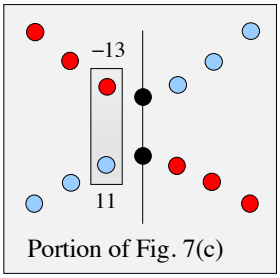
has eigenvalues and eigenfunctions:

$$E^{\pm}_{m,m'} = \tfrac{1}{2}\left(E_m + E_{m'}\right) \pm \sqrt{\tfrac{1}{4}\left(E_m - E_{m'}\right)^2 + \left|\frac{V_0}{\pi}\right|^2} \tag{1.13}$$

$$\psi^{\pm}_{m,m'} = C^{\pm}_m e^{im\phi} + C^{\pm}_{m'} e^{im'\phi}\,. \tag{1.14}$$

To obtain the coefficients $C^{\pm}_m$ and $C^{\pm}_{m'}$ requires some algebra. You should work this out and/or look it up in a quantum mechanics text.

In Fig. 7(c), $E^{\pm}_{m,m'}$ values from eqn (1.13) are shown for different m/m' combinations. Specifically, the dashed curves follow eqn (1.13), and their variation with m is recognized as an avoided crossing. Perhaps you have encountered avoided crossings in coordinate space. Here, the crossing is in "m-space," or, more generally, in k-space. Note that for $m = -m' = 12$ (black points), the eigenstates are standing waves over the low and high parts of the potential. They have no traveling wave character.



Portion of Fig. 7(c)

To visualize states such as those given by eqn (1.14), let us consider one of them: $\psi^{+}_{m,m'}$. The + superscript shall be suppressed in what follows for notational convenience, in which case we write for $\psi^{+}_{m,m'}$:

$$\psi_{m,m'} = C_m e^{im\phi} + C_{m'} e^{im'\phi}\,. \tag{1.15}$$

A little algebra yields

$$\psi_{m,m'} = \Delta C\, e^{im\phi} + 2C_{m'} e^{i(m+m')\phi/2} \cos\left(\tfrac{1}{2}(m - m')\phi\right), \tag{1.16}$$

where $\Delta C = C_m - C_{m'}$. In eqn (1.16), $\psi_{m,m'}$ is expressed in terms of traveling and standing wave character. The first term on the right hand side is a traveling wave that circulates around the ring. The second term on the right hand side has standing wave character. When $m = -m'$, it follows that $C_m = C_{m'}$, *i.e.*, $\Delta C = C_m - C_{m'}$ must be equal to zero. In this case, the first term on the right hand side of eqn (1.16) vanishes and the second term is the standing wave: $\psi_{m,m'} \propto \cos m\phi$. It is easy to calculate the probability current for the $\psi_{m,m'}$ given in eqn (1.16) by using the expression for the probability flux:

$$J = -(\hbar/m_e)\,\mathrm{Im}\left(\psi \nabla \psi^*\right). \tag{1.17}$$

This expression is derived and discussed in quantum mechanics texts.[1] Using $\nabla = \partial / r_0 \partial\phi$ (r_0 is the radius of the ring) and $\psi = \psi_{m,m'}$ yields, following some algebra,

$$J = \frac{\hbar}{m_e r_0}\left(m|C_m|^2 + m'|C_{m'}|^2 + \mathrm{Re}\left(mC_m^* C_{m'} e^{-i(m-m')\phi} + m'C_m C_{m'}^* e^{i(m-m')\phi}\right)\right). \quad (1.18)$$

Equation (18) can also be expressed in terms of velocities along the ring (v_m and $v_{m'}$) by equating the quantum mechanical angular momentum $m\hbar$ to $r_0 p_m = r_0 m_e v_m$:

$$J = v_m |C_m|^2 + v_{m'} |C_{m'}|^2 + \mathrm{Re}\left(v_m C_m^* C_{m'} e^{-i(m-m')\phi} + v_{m'} C_m C_{m'}^* e^{i(m-m')\phi}\right) \quad (1.19)$$

This confirms that J vanishes at exact resonance, where $v_m = -v_{m'}$ and $|C_m| = |C_{m'}|$. The real part is zero at resonance because the contents of the parentheses is purely imaginary. Referring to Fig. 7(c), away from resonance the currents approach v_m and $v_{m'}$ for the blue and red states, respectively. Equation (1.19) yields the flux for any eigenstate in the small V_0 regime, where the 2×2 matrix given by eqn (1.12) applies.

Threefold Symmetry

It was pointed out earlier that for an odd number of sites, the conditions for exact resonance could not be satisfied. It was stated that this would obscure the emergence of a gap for a small number like 3, whereas gaps would be obvious for systems having an odd number of sites as long as the number of sites is sufficiently large. Figure 8 verifies this.

Entries (a) and (b) for 12 and 13 sites, respectively, show gaps centered at m values that are equal to the number of sites divided by two, whether m is an integer or not. For 12 sites, the gap is at m = 6, while for 13 it is at m = 6.5. Entry (c), for 4 sites, also displays a gap. On the other hand, in (d) the blue points follow a parabola centered at the origin. Only in hindsight, and even then adding data shifted by Δm = 3 (red points), can one detect an incipient gap. Were one to peruse only the blue points, there would be no hint of a gap. Thus, in the present example the onset of behavior that can be followed to the

---

[1] The continuity equation: $\partial_t(\psi^*\psi) = (\partial_t\psi^*)\psi + \psi^*(\partial_t\psi) = -\nabla\cdot\vec{j}$, and the Schrödinger equation: $\partial_t\psi = -iH\psi$ ($\hbar$ = 1), yield: $i(\psi^* H\psi - \psi H\psi^*) = \nabla\cdot\vec{j}$.

The potential energy part of H cancels leaving: $-(i/2m)(\psi^*\nabla^2\psi - \psi\nabla^2\psi^*) = \nabla\cdot\vec{j}$. Next, the vector identity: $\nabla\cdot f\vec{g} = \nabla f\cdot\vec{g} + f\nabla\cdot\vec{g}$, yields

$$\begin{aligned}\nabla\cdot\vec{j} &= -(i/2m)\left(\nabla\cdot(\psi^*\nabla\psi) - \nabla\psi^*\cdot\nabla\psi - \nabla\cdot(\psi\nabla\psi^*) + \nabla\psi\cdot\nabla\psi^*\right)\\ &= -(i/2m)\nabla\cdot(\psi^*\nabla\psi - \psi\nabla\psi^*)\end{aligned}$$

This is the same as: $\vec{j} = (i/2m)(\psi\nabla\psi^* - \psi^*\nabla\psi)$, which is identical to eqn (1.17).

infinite lattice limit is seen with 4 sites, but not with 3 sites. If the number of segments is large, the variation will look like a smooth version of (a), with resonance at $m = N/2$.

Threefold rotational symmetry plays an important role in small molecules. It will be discussed later in Hückel theory, molecular orbital theory, and Jahn-Teller distortion.

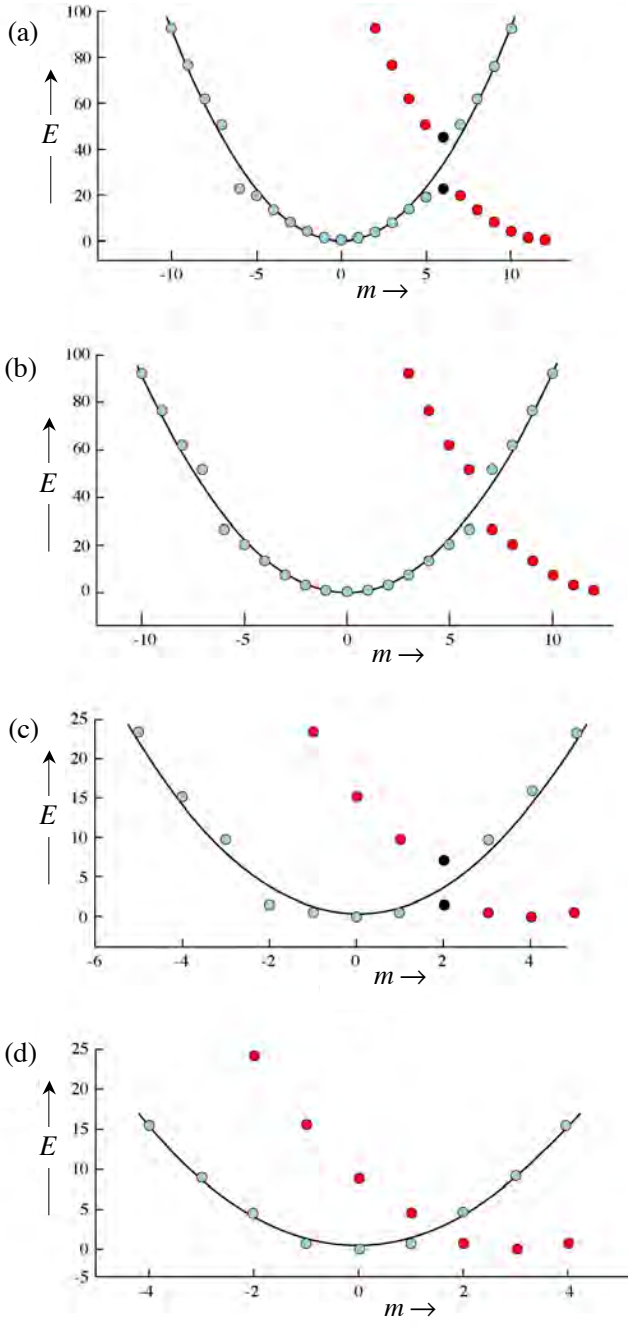


Figure 8. Plots showing energy eigenvalue versus m: Note the change in both the horizontal and vertical scales in going from the pair (a) and (b) to the pair (c) and (d). The entries (a), (b), (c), and (d) display eigenvalues for 12, 13, 4, and 3 repeating segments, respectively. The blue, red, and black points play the same roles as their counterparts in Fig. 5(c). In (a), the result shown in Fig. 5(b) has been used; a gap at $m = 6$ can be seen. In (b), it is seen that with 13 sites there is no significant difference relative to the case of 12 sites; a gap appears at $m = 6.5$. In (c), despite the small number of sites, a gap is evident at $m = 2$. In (d), the blue points fit the parabolic variation of a POR with $V(\phi) = 0$. This is due to the fact that the system is unable to satisfy simultaneously the two resonance conditions that give rise to the gaps, namely, that $|m-m'|$ is equal to an integer multiple of the number of repeating segments present in 2π, and that $m = -m'$ at the center of the resonance. This is not a big deal when there are many sites. It becomes progressively more important as the number of sites is decreased. Consequently, no gap can be seen in (d).

Less Symmetrical Potentials

Another aspect of the rectangular potential POR model involves the relationship between the length of the well and the length of each repeating segment. For example, in Fig. 5 the well width, s, is equal to one-half the repeating segment length, a. Consequently, the B_n expansion coefficients are identically zero for even n. By altering the ratio a/s, the values of B_n can be changed, including which of these are missing in the expansion. For example, with $a = 3s$, the B_n values for $n = 3, 6, 9$, etc., are zero, whereas the others are not. As a/s increases, the B_n fall off more and more gradually with n. Recall that in the delta function limit the Fourier expansion coefficients are all the same.[2]

Figure 9 shows three cases that illustrate the progression of E versus m as the ratio a/s increases. In (a), $a/s = 2$ is shown. This is the same ratio used in Figs. 7 and 8. The parameter V_0 is chosen such that there is a significant gap at $m = 6$. Thus, the lowest levels do not resemble POR states. Moreover, this value of V_0 is large enough to produce a small gap at $m = 12$, despite the fact that $B_2 = 0$. This is due to the coupling term acting twice. Matrix diagonalization reveals the effect immediately.

Figure 9(b) shows results for $a/s = 3$. The Fourier expansion of the potential gives nonzero values for B_1, B_2, and B_4 (and higher terms), whereas $B_3 = 0$. In (c), results for $a/s = 6$ are shown. Gaps appear at $m = 6, 12, 18$, and 24. It is essentially impossible to avoid small gaps at integer multiples of the first resonance (in this case at $m = 6$). This is due to the fact that at exact resonance even the smallest of perturbations gives complete mixing of the basis functions. The presence of higher-order interactions assures that *some* interaction is present. However small, it gives a 50/50 mix. In this limit, the effect cannot be seen in E versus m on the scale used in Fig. 9. It is revealed, however, through examination of the probability densities, as discussed below.

1.4. Bands of Low-Lying Levels

The last exercise in this chapter involves increasing V_0 from a small value to a significantly larger value. In this case there is an easily recognized group of lowest levels that are, to a considerable degree, localized at the sites. Wave functions of such levels have modest extension into the classically forbidden regions, and therefore tunneling rates between adjacent sites are small. These levels are nearly degenerate, and the binding sites act rather independently of one another.

When binding is weak, the electrons behave as if they were free, except near the resonances. Figure 10 shows energy eigenvalues and probability densities for 24 sites and a weak potential. In these calculations, the ratio of V_0 to the rotational constant is fixed at $V_0/B = 1$. This is sufficiently small that Fig. 10 illustrates the limit in which the electron

[2] The delta function is even about its center, so the Fourier expansion that approaches this limit uses cosines without added phase.

is, for all practical purposes, free except near the resonances. The eigenvalues follow closely a parabola (lower left), even at $m = 12$, where a strong resonance effect exists.

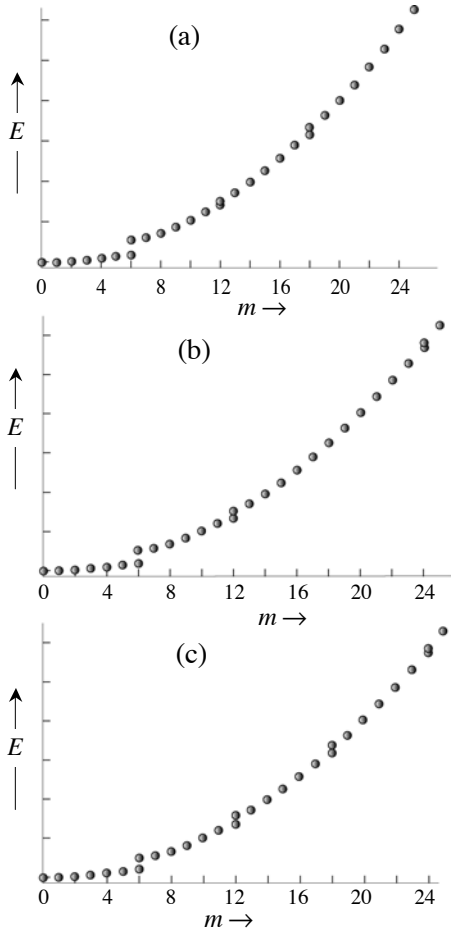


Figure 9. Energy eigenvalue versus m; the vertical scales are the same in (a), (b), and (c). Gaps are indicated with pairs of points. The energy gaps satisfy the two resonance conditions: (i) $|m' - m|$ is equal to an integer multiple of the number of repeating segments present in 2π; and (ii) $m = -m'$ at the center of the resonance. These gaps are affected by the ratio of the length of a repeating segment, a, to the width of the well, s. Results are shown for the case of 12 segments on a ring. In (a), the "square" well ($a / s = 2$) displays gaps at $m = 6$ and 18 due mainly to the first two nonzero terms in the Fourier expansion of the potential, which excludes even multiples of 12. The small gap at $m = 12$ is due to higher order interactions (see text for details). In (b), $a / s = 3$ shows gaps at $m = 6$, 12, and 24. The missing gap at $m = 18$ is due to the symmetry of this potential. In (c), $a / s = 6$ shows a gap at each integer multiple of 6. Its first missing gap would be at $m = 36$.

For $V_0 = 0$, each eigenstate has a probability density that is independent of ϕ. Referring to Fig. 10, as E increases from zero, such constancy is observed until the vicinity of the lowest resonance is encountered. This behavior is a signature of the small-V_0 regime. Of course, at exact resonance one level has $|\psi|^2$ centered over the well, while the other level has $|\psi|^2$ centered over the barrier.

Because $m = -m'$ pairs are degenerate, standing waves are formed at exact resonance with 100% efficiency because there is always some interaction present, and this is all it takes. For example, the pair $(48, -48)$ is completely coupled, whereas $m = 49$ is not coupled to any noticeable extent to its $m' = -47$ partner, because the energy difference between these basis functions is far too large to accommodate significant coupling.

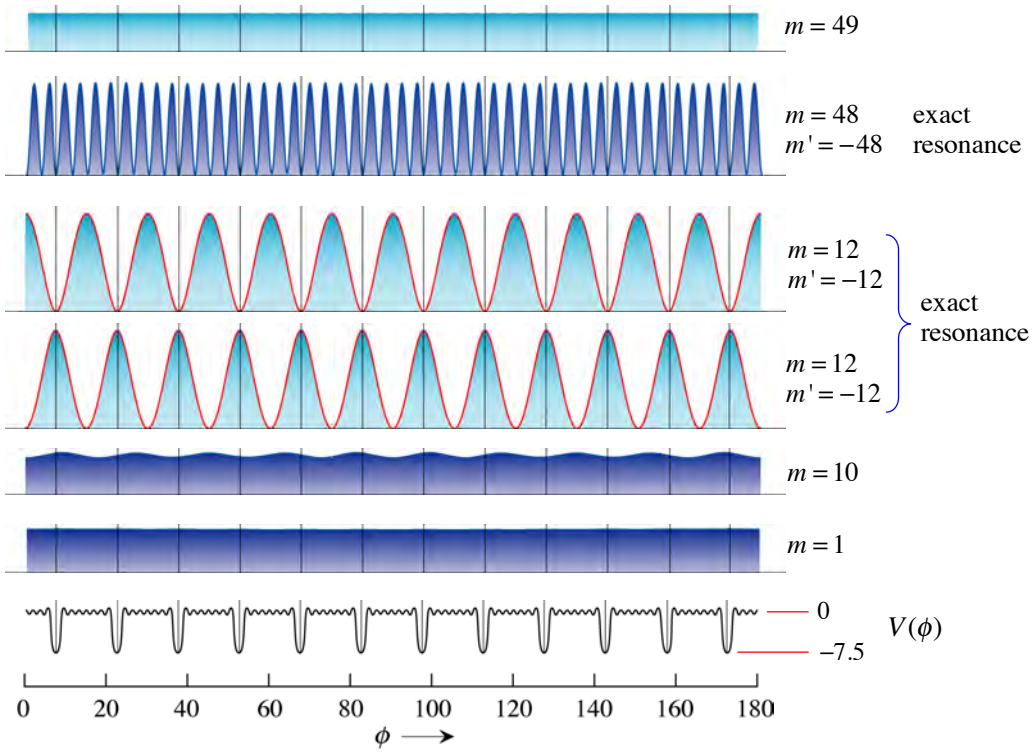


Figure 10. For small V_0, eigenvalues display near-parabolic variation (left). Probability densities (above): m = 1 is constant, while m = 10 begins to resonate. Standing waves occur at $(m, m') = (12,-12)$: the low-energy one lies over the wells; the high-energy one lies over the barriers. At (48,–48), the resonances are sharp (only one of them is shown), *i.e.*, m = 49 has constant probability density (compare this to m = 10).

Strong Binding

Now let us consider the strong binding regime. When wave functions are confined to the well regions, with only modest extension into the classically forbidden regions, we naturally think of *site functions*. Relative phases among these functions plays a big role. Though energies are affected slightly, eigenfunctions can be affected strongly because of the near degeneracy of the levels within the band. Figure 11 shows eigenvalues for 24 sites with the potential set at $V_0 = 7.5$. What is most important is the ratio of V_0 to the rotational constant B, which in this case is $V_0/B = 750$. This ratio is 750 times larger than the V_0/B ratio that was used in Fig. 10. The shape of the well is determined by the Four-

ier components. In the present case, components are chosen such that $V(\phi)$ is smoother than the rectangular well. The potential used in Figs. 10 – 12 is the one shown in Fig. 6.

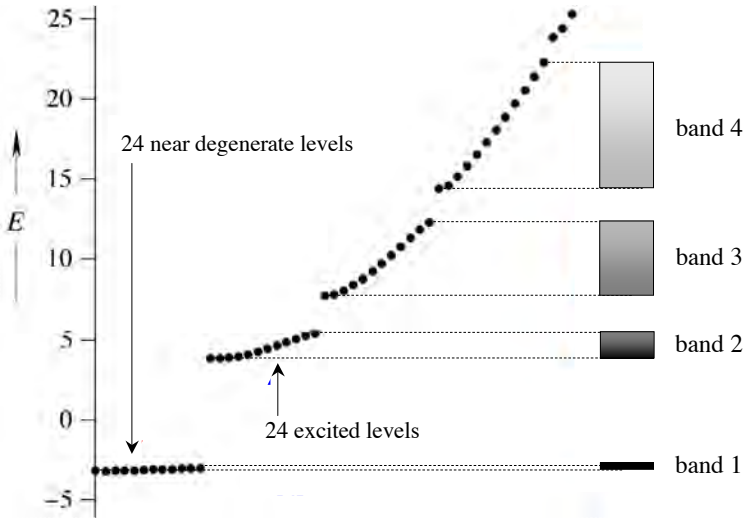


Figure 11. Parameters have been chosen such that band 1 has strong binding. Note that band 2, though narrow, has only positive energies. Band 3 is intermediate. Bands 4 and 5 (only 3 points are shown for band 5) are nearly parabolic; this is characteristic of nearly free electron behavior. The system has mirror symmetry about the vertical axis, so only half of the levels are shown: 12 rather than 24.

Figure 11 illustrates the fact that the 24 levels that constitute the lowest energy group are so strongly bound to the sites that they are nearly degenerate. There is only modest site-to-site communication. Each eigenfunction can be thought of as a linear combination of 24 localized site functions. For this lowest energy group, energies lie in a narrow band. Energy differences within this band are small compared to the gap between the lowest band and the next band. The horizontal axis simply numbers the levels. There is no relationship to the approximately good m quantum numbers used in Fig. 10. Band 2 is separated from adjacent bands by significant gaps. At higher energies, gaps persist, while at sufficiently high energy the parabolic behavior of the nearly free electron is approached.

Keep in mind that there remains a twofold degeneracy associated with clockwise and counterclockwise phase progression around the ring. This degeneracy is robust for a stationary ring and no electromagnetic interactions. One way to lift this degeneracy is through rotation of the ring itself. Another is through the $\vec{p}\cdot\vec{A}$ interaction. This arises in superconductivity and the Aharonov-Bohm effect, each of which is discussed in Part V.

The small energy differences among the levels that constitutes the lowest energy band can be understood in terms of modest penetration into the classically forbidden regions, which is inevitable because the wells are not infinitely deep. At the same time, interesting features arise because the twofold degeneracy mentioned above is robust. For example, Figs. 12 (*b*) and (*c*) are plots of $|\psi|^2$ for levels in the low and middle regions of band 1, respectively. Striking non-random undulations are seen. These are due to the computer (for whatever reason) adding and subtracting the complex degenerate functions that have constant magnitude from site to site. Presumably the complex degenerate functions are coupled through a minuscule interaction that arises in the diagonalization step.

To help visualize the strong binding case, consider 24 site functions that are identical to one another except for phase. They can serve as a basis, with eigenfunctions expressed as linear combinations of them. As mentioned above, there are small energy differences among the eigenfunctions because of differences in their kinetic and potential energies in the classically forbidden regions, that is, how the site functions overlap in the classically forbidden regions.[3]

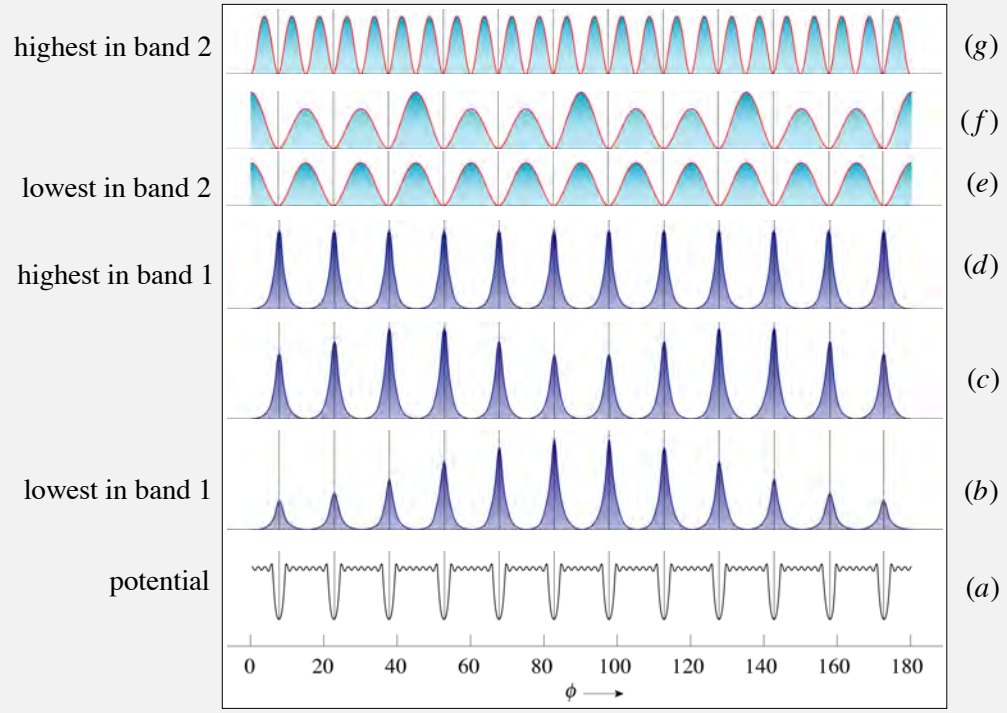


Figure 12. Probability densities for large V_0: the lowest energies correspond to localization in the well: (*b*) – (*d*) are from this group; (*d*) has the highest energy. For the next group, probability densities peak above the barriers, with minima over the wells. Entries (*e*) – (*g*) are from this group; (*e*) has the lowest energy, (*g*) the highest.

On a large scale, energy differences within the lowest band are scarcely noticeable. On a small scale, however, energy differences are evident, and the effect on the wave functions can be dramatic (Fig. 13). As mentioned above, the fact that the eigenfunctions come in degenerate pairs except at the bottom and top of the band accounts for the undulations in plots of $|\psi|^2$ versus ϕ. What has happened is that we have uncovered the irreducible representations for 24-fold rotational symmetry. Had 6 sites been used, the real orbitals of the benzene π-system might have been found.

[3] An analogous situation is inversion splitting in ammonia.

Examples from the next highest band are shown in Fig. 10(*e*)–(*g*). Note that $E > 0$ (see Fig. 11), so particles are not localized in the wells, in contrast to the lowest band. Had V_0 been assigned a sufficiently large value, there would have been multiple narrow bands with $E < 0$, and these would have been the periodic lattice analogs of the bound states of a particle in an attractive potential.

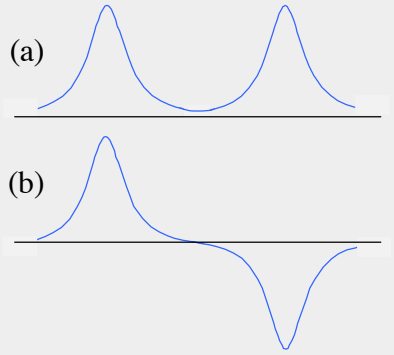


Figure 13. Probability density in the region between wells is exaggerated to illustrate overlapping site functions. (a) The site functions add in phase. (b) The site functions are 180° out of phase.

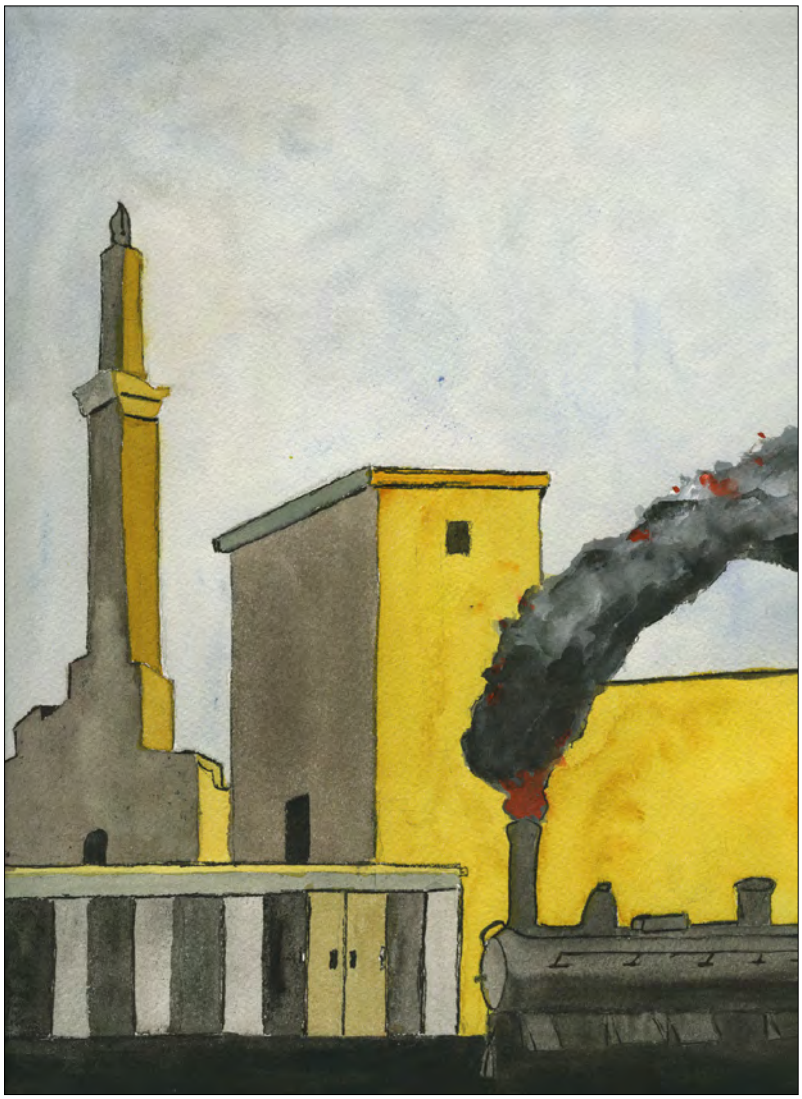
The value of V_0 used in Fig. 10 is chosen such that only one band has $E < 0$. The band centered at $E = 5$ is not analogous to a quantized state of the attractive well. It is closer in spirit to a 1D scattering resonance. In this band, the electron is biased toward being outside the region of the well. Namely, probability densities for this band have maxima in the regions between the wells, and minima at the centers of the wells. In the language of scattering theory, this band is analogous to continuum levels that lie outside the well, with minimal penetration into the well region. They do not interact with a zero order bound state to form quasibound levels that lie over the well. The next highest band: $8 < E < 12$, corresponds to a quasibound state that lies over the well, and so on for higher bands.

1.5. Summary

- For large N, allowed energies lie in bands that are separated from one another by energy gaps in which there are no levels.
- Exact resonance is defined as $m + m' = 0$, and $|m - m'| = nN$, where N is the number of sites and n labels the harmonic term that arise in the expansion of the periodic potential in a Fourier series. In terms of the wave vector k, the resonance condition yields: $k = \pm \pi/a, \pm 2\pi/a, \pm 3\pi/a \ldots$, where a is the length of a repeating unit. These regions are the edges of the Brillouin zones of a periodic lattice.
- The number of independent solutions for a band is equal to twice the number of sites when spin ½ is included.
- Electrons enter bands according to Fermi-Dirac statistics. Their transport cannot take place in a filled band. This is the case of an insulator. It can take place in an unfilled band – the case of a conductor. This transport is loss-free because of the ansatz that the binding sites are stationary (no phonons).
- Fourier components of the potential define its shape. They play a privileged role.
- Binding potentials whose magnitudes are modest relative to the POR rotational constant, result in nearly parabolic energy dependence: $E \approx Bm^2$. In this regime, the electrons are essentially free except at the band edges. The band gaps are small.
- Levels whose energies lie deep in attractive wells have narrow bands because of little site-to-site communication. As E increases, the bands become broader, reflecting increased communication between sites.
- Group velocity is zero at the edge of a band (Brillouin zone), even for nearly free electrons, because standing waves are formed by the periodic potential, however weak.
- Double degeneracy due to clockwise and counterclockwise phase progression is robust. It accounts for amusing site-to-site undulations. Adding and subtracting degenerate functions yields non-constant site-to-site probability densities. Recall the textbook pictures of benzene's molecular orbitals.

Chapter 2.

Transfer Matrices



Heading Out
Robert Wittig

Contents

Copse
Robert Wittig

Preliminary Comments

In the previous chapter, the Hamiltonian matrix was represented in the $e^{im\phi}$ basis. Its diagonalization was straightforward: (i) enter the diagonal matrix elements Bm^2; (ii) expand $V(\phi)$ in a Fourier series and enter the matrix elements of its harmonics on strips that lie parallel to the diagonal, one strip and its complex conjugate counterpart for each harmonic; and (iii) push the diagonalization button. Out comes the answer. Nothing could be easier. The useful parts of the exercise consist of varying parameters to establish regimes and limiting cases and, most importantly, interpreting the results.

The model can also be solved through exploitation of the fact that solutions must exist for which relationships between amplitudes at different (say, adjacent) sites are not site-dependent. Because the sites lie on a circle on which there is no beginning or end, these relationships depend only on the number of sites on the ring. For example, solutions must exist for which the relationship between the amplitudes at sites $i-1$ and i is identical to the relationship between the amplitudes at sites i and $i+1$. This is the basis of the approach used in this chapter.

Were non-degenerate states of our particle-on-a-ring (POR) model somehow to exist, their probability densities at different sites would have to be identical. To achieve this, the degeneracy incurred by the clockwise and counterclockwise phase progressions (differing only in sign) must be lifted. This is subtle but not unduly difficult, as seen in one of the exercises. On the other hand, what is important in the present context is the fact that for degenerate states this site-to-site constancy is not assured. In other words, degenerate states need not preserve probability density from one site to another, despite the fact that the sites are identical. After all, any linear combination of degenerate state wave functions is a solution of the Schrödinger equation, so it comes as no surprise that degenerate state probability densities are not necessarily the same at different sites.[1]

2.1. The Model

The above caveats about degenerate states notwithstanding, for the case of a periodic potential on a ring, it is always possible to obtain wave functions that preserve probability density in going from one site to the next. Moreover, in the absence of an external field, any state that has a circulating flux must have a degenerate counterpart whose flux has the same magnitude but circulates in the opposite direction.

Referring to Fig. 1, which depicts a case of fairly strong binding, $\Phi_i^n(x)$ shall be referred to as a site amplitude. The wave function $\psi^n(x)$ is an eigenfunction of the Schrödinger equation for the periodic potential $V(x)$. Therefore, $\psi^n(x)$ extends over all sites,

---

[1] If these states are filled with particles (electrons), the site-to-site constancy is recovered. Recall the π system of benzene. Its HOMO is a degenerate pair that is often expressed as real functions that do not preserve probability density between sites. But when six electrons are put in the molecular orbitals, the charge density is the same at each carbon atom.

whereas $\Phi_i^n(x)$ is local to site i. Note that location on the ring is given by $x = r_0\phi$, where r_0 is the radius of the ring. Either x or ϕ can be used.

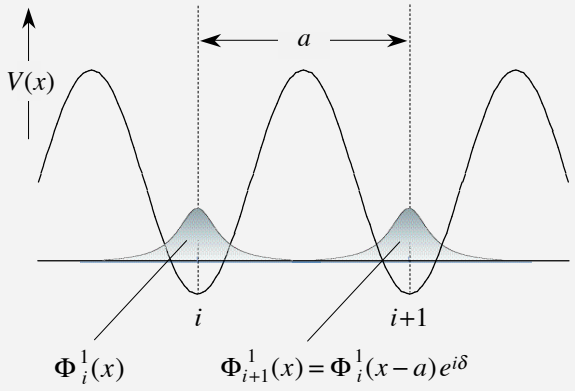


Figure 1. The function centered at site $i+1$ is related to the function centered at site i by the phase factor $e^{i\delta}$. Different values of phase shift δ correspond to slightly different energies when there is modest overlap between functions centered at adjacent sites. The lowest (n = 1) site amplitudes are shown. The superscript 1 denotes n = 1. Were overlap between $\Phi_i^1(x)$ and $\Phi_{i+1}^1(x)$ zero, N such sites would be N-fold degenerate.

The index n that appears in $\Phi_i^n(x)$ and in $\psi^n(x)$ reflects the site character. For example, if binding is strong enough to localize electrons at the sites, n is counterpart to the quantum number for a bound level of a site. This strong binding regime applies to the n = 1 case indicated in Fig. 1. This case is analogous to a low-lying atomic level that is affected only slightly by nearby atoms and therefore presents a narrow band. However, the condition we are imposing on the site functions is general, applying to strong binding, weak binding, and everything in-between. Thus, we require that

$$\Phi_{i+1}^n(x) = \Phi_i^n(x-a)e^{i\delta}, \tag{2.1}$$

where δ is the phase shift between adjacent sites. Though degenerate eigenfunctions need not preserve probability density from site to site, we will stick with eqn (2.1). Namely, site functions are chosen such that all site-to-site variation is accounted for with a phase factor $e^{i\delta}$. Such solutions must exist.

Boundary Condition

The wave function is required to be single valued everywhere on the ring. Thus, when it is complex – which is necessary if there is to be a circulating flux – its phase cannot be discontinuous (mod 2π) at any point.[2] If there are N sites, the round-trip phase shift, $N\delta$, must be $2\pi k$, where k is an integer. In this case the site-to-site phase shift is

[2] Discontinuous phase arises when a geometric phase is present. Dealing with this requires that a gauge field accompany the canonical momentum. This has no significance in the present context. Also, when spin is included, the system need not recover its original phase upon rotation by 2π. For example, a spin ½ ket changes sign when it is rotated by 2π, whereas it is single valued in 4π.

$$\delta = \frac{2\pi}{N} k \,. \tag{2.2}$$

The integer k that enters through the boundary condition on the round-trip phase is a quantum number. We shall see that it is closely related to the wave vector (also labeled k) that was introduced in Chapter 1. Hopefully, the fact that the symbol k is used for each will not lead to confusion. Note that the quantum number is dimensionless, whereas the wave vector has dimension of inverse length.

Equation (2.2) applies separately to each value of n. For example, Fig. 1 depicts $n = 1$, but the same rule applies to other n levels (Fig. 2). As mentioned earlier, when binding is strong enough that an electron is highly localized at a site, the index n is analogous to the quantum number for a low-lying atomic level.

On the other hand, in the weak-binding limit the lattice sites are less important, as quantization is dominated by the round-trip distance, $Na = L$. In this limit, we are concerned mainly with the round-trip phase factor: $e^{iN\delta} = e^{i2\pi k}$, which is the same as that for unencumbered ($V(\phi) = 0$) POR waves. The discrete nature of the lattice is only important near the resonances that are present when integer multiples of half wavelengths are comparable to the lattice spacing, a (recall Fig. 7 in Chapter 1).

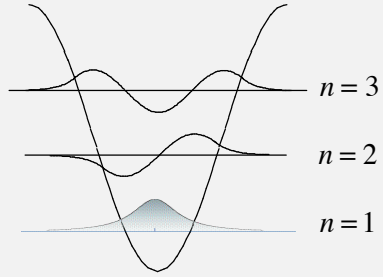


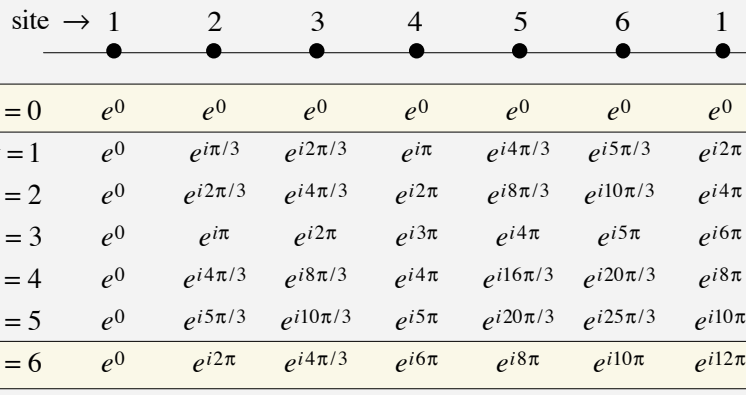
Figure 2. The index n labels excited states of a single segment.

Allowed *k* Values

Let us now establish the allowed k values. Namely, how many independent k values exist, and which integers constitute this group. This can be inferred from Figs. 1 and 2, which depict the strong-binding limit. For a given n value, a basis of N localized orbitals gives rise to N eigenstates, leaving aside spin. Therefore, this is the number of independent k values for each n value. This number has nothing to do with the strength of the potential. Once these solutions are in hand, if additional k values are introduced, no additional independent states are forthcoming. States are merely repeated.

To see how this works, consider k values that range from 0 to $N-1$. For $k = 0$ we have, from eqns (2.1) and (2.2), $\Phi_{i+1}^{n}(x) = \Phi_{i}^{n}(x-a)$. That is, the site-to-site phase shift δ is zero. The other extreme is $k = N-1$, where $\delta = (N-1)2\pi/N$. In this case, the round-trip phase shift is $N\delta = (N-1)2\pi$. Note that δ never reaches 2π. When N is large this is unimportant, and for all practical purposes δ can be taken as reaching 2π. However, for small N the fact that δ does not reach 2π is important. For k values between 0 and $N-1$, δ increases linearly with k, with each k value corresponding to a distinct state. Thus, there are N distinct states, one for each value of k.

Now let us see what happens if k is allowed to take on the value N instead of stopping at $N-1$. In this case $\delta = 2\pi$. However, a phase shift of 2π is equivalent to a phase shift of zero. Thus, there is no difference between the states $k = 0$ and $k = N$. Likewise, if we let $k = N+1, N+2, N+3\ldots$, we see that this repeats the states associated with $k = 1, 2, 3\ldots$. This explains why, for a given n, only N values of k can be associated with independent states. Figure 3 illustrates this for $N = 6$.



$k=0$	e^{0}	e^{0}	e^{0}	e^{0}	e^{0}	e^{0}	e^{0}
$k=1$	e^{0}	$e^{i\pi/3}$	$e^{i2\pi/3}$	$e^{i\pi}$	$e^{i4\pi/3}$	$e^{i5\pi/3}$	$e^{i2\pi}$
$k=2$	e^{0}	$e^{i2\pi/3}$	$e^{i4\pi/3}$	$e^{i2\pi}$	$e^{i8\pi/3}$	$e^{i10\pi/3}$	$e^{i4\pi}$
$k=3$	e^{0}	$e^{i\pi}$	$e^{i2\pi}$	$e^{i3\pi}$	$e^{i4\pi}$	$e^{i5\pi}$	$e^{i6\pi}$
$k=4$	e^{0}	$e^{i4\pi/3}$	$e^{i8\pi/3}$	$e^{i4\pi}$	$e^{i16\pi/3}$	$e^{i20\pi/3}$	$e^{i8\pi}$
$k=5$	e^{0}	$e^{i5\pi/3}$	$e^{i10\pi/3}$	$e^{i5\pi}$	$e^{i20\pi/3}$	$e^{i25\pi/3}$	$e^{i10\pi}$
$k=6$	e^{0}	$e^{i2\pi}$	$e^{i4\pi/3}$	$e^{i6\pi}$	$e^{i8\pi}$	$e^{i10\pi}$	$e^{i12\pi}$

Figure 3. Phase progression for different k: The site 1 phase is set to zero. For $k = 0-5$, the states are distinct. For $k = 6$, the phases at the sites are the same as for $k = 0$, so $k = 0$ and 6 represent the same state. The $k = 7$ and $k = 1$ states are also the same, and so on.

A pictorial way to illustrate this is given in Fig. 4. Adding an integer number of cycles between adjacent sites does not affect the site phases. In the upper two entries, $2\pi/a$ and $4\pi/a$ have been added to k (blue) to obtain the red curves. This adds one and two cycles, respectively, between the black points. In the lowest entry one cycle is added to a lattice that is twice as dense.

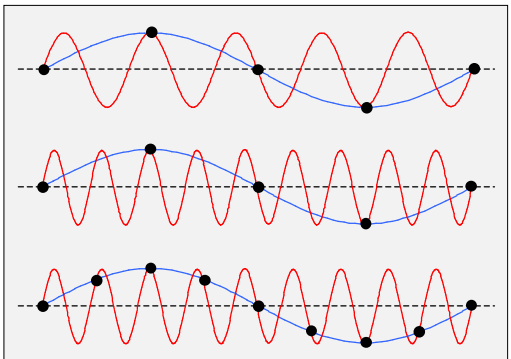


Figure 4. Site phase is unaffected when k is increased by an integer multiple of $2\pi/a$.

In the above example, k values of 0, 1, 2…$N-1$ have been used to make the point. However, the range *per se* does not matter as long as it contains (or can be made to contain) N successive integers.[3] Symmetric ranges are the most common. For odd N, this spans $-(N-1)/2$ to $(N-1)/2$, whereas even N spans $-N/2$ to $N/2$, with the understanding that the two end points, taken together, contribute one state, not two. The use of both $-N/2$ and $N/2$ would

[3] Amusing collections of k values can be organized into a sequence of successive integers by adding or subtracting multiples of N. For example, consider $N = 6$ and k values 0, 1, 8, 3, 10, and 17. This is equivalent to 0, 1, 2, 3, 4, and 5 because 8, 10, and 17 differ from 2, 4, and 5 by the addition of multiples of 6. N allowed values of k can always be expressed as successive integers.

over-count the number of independent states by one. In other words, a linear combination of $-N/2$ and $N/2$ states can be taken, as long as only one of the linear combinations is used. This is like the black points used in Fig. 7 of Chapter 1.

There are many ways to obtain the solutions. The one used in the present chapter expresses the eigenfunctions in terms of the relationship between amplitudes at adjacent sites. The Schrödinger equation is a second order differential equation. For unimpeded motion, there are two independent general solutions for each value of $|k|$ that, upon applying the boundary conditions, give the eigenvalues and eigenfunctions. Because the potential is periodic, it is not difficult to cast the eigenvalue problem in terms of an individual site.

Transfer Matrices

In the vicinity of one of the sites we need not, at this time, specify the two independent solutions. It is sufficient to acknowledge that two amplitudes are needed. For example, the relationship between the amplitudes at sites i and $i+1$ can be expressed as

$$\begin{pmatrix} A_{i+1} \\ B_{i+1} \end{pmatrix} = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \begin{pmatrix} A_i \\ B_i \end{pmatrix}. \tag{2.3}$$

Figure 5 shows A and B as coefficients of plane waves. The use of plane waves is convenient. For transit through a single region of potential the plateaus on the two sides can have different heights. However, for the closed (circular) systems under consideration, the plateaus must be of the same height.

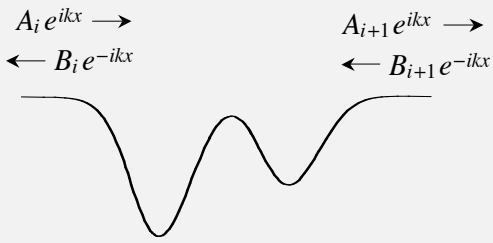


Figure 5. Scattering via the potential $V(x)$ transforms incoming waves to outgoing waves.

The site-to-site transfer of amplitude throughout a closed periodic lattice can be modeled using the T-matrix in eqn (2.3) to progress from one site to the next until we arrive back at the original site. The net effect is unity, as the final amplitudes are, by definition, equal to the original amplitudes. For N identical sites, matrix multiplication gives

$$T^N = 1. \tag{2.4}$$

The use of matrix multiplication to transfer amplitude has been around a long time. When the author was an undergraduate, it was used to analyze transmission lines and electronic circuits. Leon Brillouin made frequent use of it in his 1946 book [14]: *Wave Propagation in Periodic Structures*. I found a few books that discuss its use in quantum mechanics and they are listed in the bibliography at the end of Part A. The book by Gilmore

[16] is straightforward and I will follow several of the arguments given there. The book by Levi [15] also covers this topic, albeit with typos and small errors.

The idea is to use eqn (2.3) together with the fact that the amplitudes vary from site to site only by a phase shift:

$$\begin{pmatrix} A_{i+1} \\ B_{i+1} \end{pmatrix} = e^{i\delta} \begin{pmatrix} A_i \\ B_i \end{pmatrix}. \tag{2.5}$$

This is one of the most important equations in this chapter. It is the embodiment of the preceding discussion and the basis of the math that follows. Combining eqns (2.3) and (2.5) gives

$$\begin{pmatrix} T_{11}-\lambda & T_{12} \\ T_{21} & T_{22}-\lambda \end{pmatrix} \begin{pmatrix} A_i \\ B_i \end{pmatrix} = 0 \,, \tag{2.6}$$

where $\lambda = e^{i\delta}$ is the eigenvalue. The eigenvalues of eqn (2.6) are

$$\lambda = \tfrac{1}{2}\mathrm{Tr}\,T \pm \underbrace{\tfrac{1}{2}\sqrt{T_{11}{}^2 + T_{22}{}^2 - 2T_{11}T_{22} + 4T_{12}T_{21}}}_{\left(\left(\frac{1}{2}\mathrm{Tr}\,T\right)^2 - \det T\right)^{1/2}} \,, \tag{2.7}$$

where Tr denotes trace. Using the fact that $\det T = 1$ yields

$$\lambda = \tfrac{1}{2}\mathrm{Tr}\,T \pm i\left(1 - \left(\tfrac{1}{2}\mathrm{Tr}\,T\right)^2\right)^{1/2}. \tag{2.8}$$

$$= e^{i\delta} = \underbrace{\cos\delta}_{\frac{1}{2}\mathrm{Tr}\,T} + \underbrace{i\sin\delta}_{\pm i\left(1-\left(\frac{1}{2}\mathrm{Tr}\,T\right)^2\right)^{1/2}} \tag{2.9}$$

All of the information in eqn (2.9) is contained in eqn (2.10) below, because if we know the real part of λ its imaginary part is automatically known. Thus, using $\delta = 2\pi k/N$, eqn (2.9) yields

$$\cos\left(\frac{2\pi}{N}k\right) = \tfrac{1}{2}\mathrm{Tr}\,T \,. \tag{2.10}$$

As mentioned earlier, the parameter k that enters through the boundary condition on the round-trip phase is a quantum number. Recall that in Chapter 1 matrices were diagonalized using the $e^{im\phi}$ basis. The eigenstates were linear combinations of the basis functions, with m being a good quantum number only in the limit $V(\phi) \to 0$. We now see that

eigenstates can always be labeled with the quantum number k. In the limit $V(\phi) \to 0$, k and m are the same, whereas k is a good quantum number for all values of $V(\phi)$.

This can be stated more elegantly. The quantum number m is the result of a continuous symmetry, with angular momentum being its conserved flux. The quantum number k is the discrete symmetry counterpart. Though it is a good quantum number for the discrete symmetry of the periodic potential, it has no conserved flux because the symmetry is not continuous. Only a continuous symmetry can have a conserved flux associated with it. This is an application of an important theorem introduced by the mathematician Emmy Noether.

As promised, the model is solved using a single site. The number of sites enters only trivially in the argument of the cosine in eqn (2.10). This transcendental equation accounts for the entire set of eigenvalues and eigenfunctions: all n values and all k values. What is needed now is the transfer matrix T.

2.2. Rectangular well

The rectangular well that was introduced in Chapter 1 is given here as Fig. 6. The strategy is to use a single repeating element, and write the T matrix that relates waves in this element to those in an adjacent element. This enables eqn (2.10) to be used. The easiest way to carry out the math is with an algorithm for transferring amplitude through regions of piecewise constant potential. This approach can be used to model 1D potentials having any shape, because they can always be approximated using a series of potential steps. Details are given in the Appendix at the end of this chapter. It turns out that the math is most straightforward if, instead of the T matrix, a matrix M is used that is similar to T, but easier to work with. Referring to Fig. 6, matrices M_i transfer the system *between* boundaries, whereas matrices T_i transfer the system *across* boundaries. Equation (2.10) then applies with M replacing T. The reader is encouraged very strongly indeed to read the Appendix before proceeding.

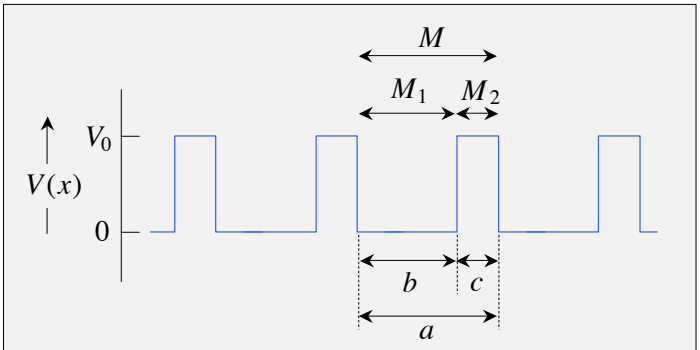


Figure 6. The well length is b, the barrier length is c, and the length of the unit cell is $a = b + c$. The matrices M, M_1, and M_2 are discussed in the text and in the Appendix.

Transcendental

For the unit cell $a = b + c$ in Fig. 6, M is obtained by multiplication: $M = M_1 M_2$, where M_1 and M_2 transform amplitude over distances b and c, respectively. Using results from the Appendix, for $E < V_0$, the matrices M_1 and M_2 are

$$M_1 = \begin{bmatrix} \cos Kb & -K^{-1}\sin Kb \\ K\sin Kb & \cos Kb \end{bmatrix} \qquad M_2 = \begin{bmatrix} \cosh \beta c & -\beta^{-1}\sinh \beta c \\ -\beta \sinh \beta c & \cosh \beta c \end{bmatrix} \tag{2.11}$$

where

$$\hbar K = \left(2mE\right)^{1/2} \qquad \hbar\beta = \left(2m(V_0 - E)\right)^{1/2} . \tag{2.12}$$

Alternatively, for $E > V_0$, β is imaginary, and the hyperbolic functions in M_2 become sines and cosines. In this case, the mathematical form of M_2 is the same as that of M_1, but with wave vector $[2m(E-V_0)/\hbar^2]^{1/2}$ instead of $(2mE/\hbar^2)^{1/2}$. To obtain the desired expression for $E < V_0$, eqn (2.10) is used with T replaced by $M = M_1M_2$. Carrying out the matrix multiplication and taking one half the trace gives

$$\cos\left(\frac{2\pi}{N}k\right) = \cos Kb\, \cosh \beta c + \frac{1}{2}\left(\frac{\beta}{K} - \frac{K}{\beta}\right)\sin Kb\, \sinh \beta c . \tag{2.13}$$

This transcendental equation defines the allowed energies. To see how it works, increase E, starting from zero, and put the resulting K and β values into eqn (2.13). This enables one to see which values of E are allowed, namely, which values of E satisfy eqn (2.13) and therefore correspond to eigenvalues. For example, when the magnitude of the right hand side of eqn (2.13) exceeds unity the corresponding E values are not allowed because the magnitude of the left hand side is always less than or equal to unity. In physical terms, when the barriers between binding sites are sufficiently large to ensure localization near the sites, there can be no solutions for which E is ~ 0. The boundary conditions at the barriers simply impose too much curvature on the wave function, and therefore too much kinetic energy, to accommodate solutions having $E \sim 0$. As E is increased from zero, the right hand side of eqn (2.13) decreases until it reaches a value of $+1$. This is the first allowed energy. It corresponds to $k = 0$.

Figure 7 illustrates eqn (2.13) for the case $N = 6$. Such a graph enables us to identify all of the eigenvalues: the energies corresponding to each of the k values, for each of the n values. Let us follow this progression along, step-by-step, starting at $E = 0$.

To begin, consider the red circle that has unit radius. The left hand side of eqn (2.13) ($\cos(2\pi k/N) = \cos\delta$) is equal to $\cos(\pi k/3)$ for $N = 6$. It is represented by projections of the unit radius vector on the vertical axis for the allowed k values. For example, $k = 0$ is the point at the top of the circle. Horizontal dashed lines indicate the values of $\cos\delta = \cos(\pi k/3)$ for the allowed k values.

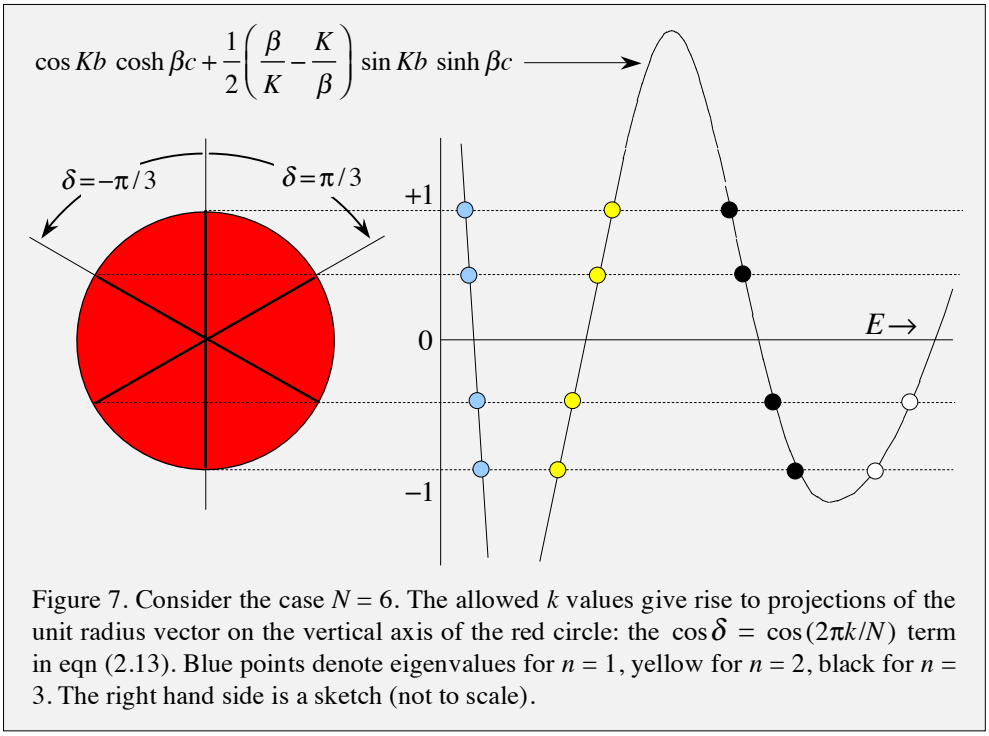


Figure 7. Consider the case $N = 6$. The allowed k values give rise to projections of the unit radius vector on the vertical axis of the red circle: the $\cos\delta = \cos(2\pi k/N)$ term in eqn (2.13). Blue points denote eigenvalues for $n = 1$, yellow for $n = 2$, black for $n = 3$. The right hand side is a sketch (not to scale).

Now go to the right hand side of eqn (2.13). With $E = 0$, we get a number that exceeds unity, for which there is no solution. Increase E until the right hand side is + 1. This is $k = 0$. It corresponds to the blue point having the smallest E. The next blue point corresponds to $k = \pm 1$ ($\delta = \pm\pi/3$), a double degeneracy. Next we have $k = \pm 2$ ($\delta = \pm 2\pi/3$), again a double degeneracy. The last blue point corresponds to $k = \pm 3$ ($\delta = \pm\pi$); however, only one state is retained. The four blue points correspond to the six $n = 1$ states.

2.3. Degenerate Pairs

Now consider some consequences of the fact that all of the $n = 1$ (blue) states except those at the top and bottom of the red circle in Fig. 7 come in doubly degenerate pairs. This is true regardless of the value of N. When N is even there is a point at the bottom of the red circle. When N is odd there is not.

Recall that when the POR model was diagonalized in the strong binding regime, interesting wave functions were revealed. For example, in Fig. 12 of Chapter 1, entries (*b*), (*c*), and (*f*) show probability densities that vary from site to site, even though the potential is identical from site to site. In fact, though it was not shown in Fig. 12 of Chapter 1, in some cases the probability density goes to zero at certain sites.

Because the sites are identical, we know that solutions *must* exist for which probability densities are unchanged when x is displaced by an integer multiple of the spacing between adjacent sites. Moreover, the probability densities shown in Fig. 12 (*b*), (*c*), and (*f*)

of Chapter 1 must correspond to wave functions that are combinations of ones that differ only in the sense of phase progression: $e^{i2\pi k/N}$ and $e^{-i2\pi k/N}$.

It is now possible to interpret these results more quantitatively. When a computer program diagonalizes a matrix, it can introduce minuscule energy differences in levels that are, in fact, rigorously degenerate because of symmetry. These differences appear in some distant decimal place. By themselves, they are of no concern. However, even the slightest perturbation causes the wave functions to be 50/50 mixtures of the degenerate functions. Thus, the appearance of real functions whose magnitudes vary from site to site should come as no surprise. It is a manifestation of the underlying degeneracy.

To illustrate this effect, consider $N = 6$ and the phase progression given by eqns (2.1) and (2.2) with $k = \pm 1$ ($\delta = \pm \pi/3$). The wave functions for $k = +1$ and -1 are now written using eqns (2.1) and (2.2):

$$\psi(+1) = \chi_1 + \chi_2 e^{i\pi/3} + \chi_3 e^{i2\pi/3} + \chi_4 e^{i\pi} + \chi_5 e^{i4\pi/3} + \chi_6 e^{i5\pi/3} \tag{2.14}$$

$$\psi(-1) = \chi_1 + \chi_2 e^{-i\pi/3} + \chi_3 e^{-i2\pi/3} + \chi_4 e^{-i\pi} + \chi_5 e^{-i4\pi/3} + \chi_6 e^{-i5\pi/3} \tag{2.15}$$

where χ_i is a localized (real) orbital at the i^{th} site. Normalization is suppressed. It can be carried out later should the need arise. The addition and subtraction of eqns (2.14) and (2.15) gives

$$\psi(+1) + \psi(-1) = 2\chi_1 + \chi_2 - \chi_3 - 2\chi_4 - \chi_5 + \chi_6 \tag{2.16}$$

$$\psi(+1) - \psi(-1) = \chi_2 + \chi_3 - \chi_5 - \chi_6 \tag{2.17}$$

Equations (2.16) and (2.17) illustrate the fact that combinations of the degenerate pair account for the probability density varying from site to site, even going to zero at some sites. The spatial frequency of this variation is proportional to k. When k is zero, the system is non-degenerate and the site functions are all in phase and have the same magnitudes at each of the sites. This is a bonding situation. When $|k| = 3$, the system is also non-degenerate. The site functions have the same magnitudes at each of the sites but they have alternating signs in going from site to site. This is an antibonding situation. As anticipated, we have simply reproduced some well-known results for the benzene molecule.

Higher n

Continuing with Fig. 7, as E is increased past the lowest blue point, there are at first no solutions because the curve's value is less than -1. However, the curve eventually turns around and increases until it again reaches the value -1. This is the lowest yellow point. This eigenvalue corresponds to $n = 2$. As E continues to increase, we find two pairs of doubly degenerate states and a non-degenerate state at $k = 0$. These yellow points constitute the eigenvalues for the six $n = 2$ states. As E continues to increase, we continue to get groups of 6 points (black, white, and so on) corresponding to $n = 3$, 4, and so on.

When E gets large enough, the particle is essentially free to move around the ring. It then experiences little localization. In this limit, the wave vector K will be quantized according to the overall length of the ring, $L = N(b+c) = Na$, with k serving as the quantum number. This limit is easily obtained from eqn (2.13). With the substitution $\beta \to iK$, which is appropriate to the limit $V_0 \to 0$, the right hand side of eqn (2.13) becomes

$$\cos Kb \cos Kc - \sin Kb \sin Kc . \tag{2.18}$$

But this is simply $\cos K(b+c) = \cos Ka$. Comparing the left and right hand sides of eqn (2.13), we see that

$$\frac{2\pi}{N}k = Ka \implies K = \frac{2\pi}{L}k , \tag{2.19}$$

where $L = Na$. We have simply recovered the POR result for $V(\phi) = 0$.

2.4. Summary

- Solutions for the 1D periodic potential are obtained by requiring that site amplitudes satisfy $\Phi_{i+1}^{n}(x) = \Phi_{i}^{n}(x-a)e^{i\delta}$, where δ is the phase shift between adjacent sites. The fact that the round-trip phase shift must be equal to an integer multiple of 2π gives a site-to-site phase shift of $\delta = 2\pi k/N$, where k is a quantum number.
- Because $T^N = M^N = 1$, the computational task is reduced to that of a single site. It does not depend on the number of sites, except trivially.
- The use of transfer matrices (either T or M) enables the method to be applied to potentials of any shape using a simple algorithm.
- Solutions are defined by the transcendental equation: $\cos(2\pi k/N) = \frac{1}{2}\mathrm{Tr}\,T$.
- A graphical mnemonic enables one to identify features such as bands of levels, including widths, degeneracy, gaps, density of states, and strong and weak binding limits.

Appendix: Transfer Matrices

It is often useful to approximate a smooth 1D potential $V(x)$ by joining together, at step discontinuities, regions of constant potential, as indicated in Fig. 1. Situations in which this proves expedient arise in many contexts: transmission through a potential barrier; bound states; quasibound states (resonances); periodic potentials, including infinite lattices; and so on. In this appendix, piecewise constant approximations of $V(x)$, such as the one shown in Fig. 1, are used to construct an algorithm. The fact that $\psi(x)$ and its derivative are continuous at a step discontinuity yields the same mathematical relationship between wave amplitudes on either side of the boundary for all boundaries.

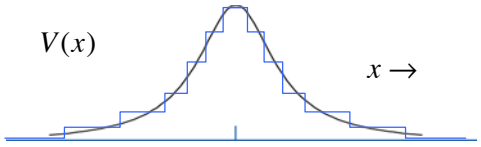


Figure 1. The smooth potential (black) can be approximated by a series of steps (blue), for any well behaved $V(x)$.

Referring to Fig. 2, the equations that equate the amplitudes and momenta on either side of the step boundary separating regions 1 and 2 are

$$\begin{pmatrix} e^{ik_1a_1} & e^{-ik_1a_1} \\ ik_1e^{ik_1a_1} & -ik_1e^{-ik_1a_1} \end{pmatrix} \begin{pmatrix} A_1 \\ B_1 \end{pmatrix} = \begin{pmatrix} e^{ik_2a_1} & e^{-ik_2a_1} \\ ik_2e^{ik_2a_1} & -ik_2e^{-ik_2a_1} \end{pmatrix} \begin{pmatrix} A_2 \\ B_2 \end{pmatrix}, \tag{1}$$

where $\hbar k_i = [2m(E - V_i)]^{1/2}$.

Figure 2 and eqn (1) are expressed in a way that is appropriate to the regime $E > V_2$, in which case k_2 is real. Alternatively, for $E < V_2$, k_2 is imaginary. In this case, $k_2 \to i\beta_2$, where $\hbar\beta_2 = [2m(V_2 - E)]^{1/2}$. The math is the same as for the case of real k_2 (albeit with $e^{ik_2x} \to e^{-\beta_2 x}$) even though the physics is quite different: propagation versus attenuation.

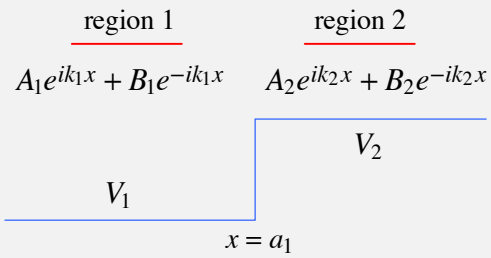


Figure 2. Waves present in regions 1 and 2 are matched at the $x = a_1$ boundary. Likewise, momentum is continuous at a_1.

Equation (1) is easily extended to include additional steps. Indeed, the multiplication of a string of matrices enables us to deal with any shape that can be approximated by using a series of steps (Fig. 1). Sometimes it is useful to express matrices such as those in eqn (1) in terms of even simpler matrices:

$$\begin{pmatrix} e^{ika} & e^{-ika} \\ ike^{ika} & -ike^{-ika} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ ik & -ik \end{pmatrix} \begin{pmatrix} e^{ika} & 0 \\ 0 & e^{-ika} \end{pmatrix}. \tag{2}$$

Using this with eqn (1) gives

$$\begin{pmatrix} A_1 \\ B_1 \end{pmatrix} = \begin{pmatrix} e^{ik_1a_1} & 0 \\ 0 & e^{-ik_1a_1} \end{pmatrix}^{-1} \begin{pmatrix} 1 & 1 \\ ik_1 & -ik_1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 1 \\ ik_2 & -ik_2 \end{pmatrix} \begin{pmatrix} e^{ik_2a_1} & 0 \\ 0 & e^{-ik_2a_1} \end{pmatrix} \begin{pmatrix} A_2 \\ B_2 \end{pmatrix}. \tag{3}$$

Evaluating the two inverse matrices yields

$$\begin{pmatrix} A_1 \\ B_1 \end{pmatrix} = \begin{pmatrix} e^{-ik_1a_1} & 0 \\ 0 & e^{ik_1a_1} \end{pmatrix} \frac{1}{2} \begin{pmatrix} 1 & -i/k_1 \\ 1 & i/k_1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ ik_2 & -ik_2 \end{pmatrix} \begin{pmatrix} e^{ik_2a_1} & 0 \\ 0 & e^{-ik_2a_1} \end{pmatrix} \begin{pmatrix} A_2 \\ B_2 \end{pmatrix} \tag{4}$$

$$= T_{12} \begin{pmatrix} A_2 \\ B_2 \end{pmatrix}. \tag{5}$$

The matrix T_{12} transfers amplitude across the boundary separating regions 1 and 2, hence the term transfer matrix. Moreover, eqn (5) is extended to n successive regions by compounding eqn (5):

$$\begin{pmatrix} A_1 \\ B_1 \end{pmatrix} = T_{12}\, T_{23}\, \ldots\, T_{n\text{-}1,n} \begin{pmatrix} A_n \\ B_n \end{pmatrix}. \tag{6}$$

The above manipulations show that by approximating $V(x)$ with a piecewise constant potential the transmission problem is reduced to the multiplication of a string of 2×2 matrices. This is easily carried out using Mathcad or some other software package.

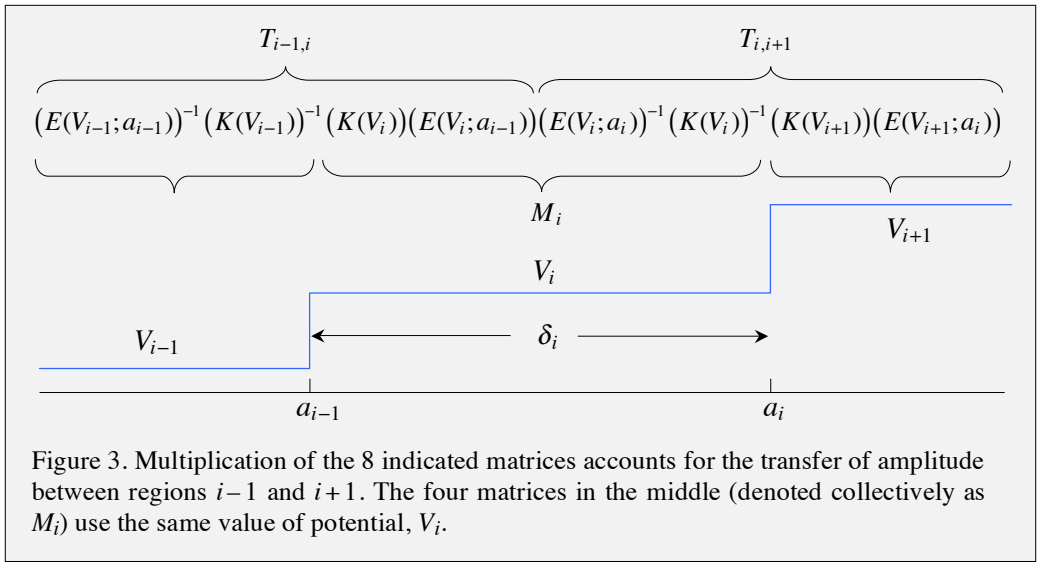


Figure 3. Multiplication of the 8 indicated matrices accounts for the transfer of amplitude between regions $i-1$ and $i+1$. The four matrices in the middle (denoted collectively as M_i) use the same value of potential, V_i.

Further simplification, insofar as the computer algorithm is concerned, can be obtained through a judicious grouping of the matrices. For example, Fig. 3 illustrates the net transmission through the step boundaries at a_{i-1} and a_i (going from $x < a_{i-1}$ to $x > a_i$).

Note that a shorthand notation for the matrices has been introduced. The matrices indicated in Fig. 3 are defined as follows:

$$
\begin{aligned}
\left(E(V_{i-1};a_{i-1})\right)^{-1} &= \begin{pmatrix} e^{-ik_{i-1}a_{i-1}} & 0 \\ 0 & e^{ik_{i-1}a_{i-1}} \end{pmatrix} & \left(K(V_{i-1})\right)^{-1} &= \frac{1}{2}\begin{pmatrix} 1 & -i/k_{i-1} \\ 1 & i/k_{i-1} \end{pmatrix} \\
\left(E(V_i;a_{i-1})\right) &= \begin{pmatrix} e^{ik_i a_{i-1}} & 0 \\ 0 & e^{-ik_i a_{i-1}} \end{pmatrix} & \left(K(V_i)\right) &= \begin{pmatrix} 1 & 1 \\ ik_i & -ik_i \end{pmatrix} \\
\left(E(V_i;a_i)\right)^{-1} &= \begin{pmatrix} e^{-ik_i a_i} & 0 \\ 0 & e^{ik_i a_i} \end{pmatrix} & \left(K(V_i)\right)^{-1} &= \frac{1}{2}\begin{pmatrix} 1 & -i/k_i \\ 1 & i/k_i \end{pmatrix} \\
\left(E(V_{i+1};a_i)\right) &= \begin{pmatrix} e^{ik_{i+1}a_i} & 0 \\ 0 & e^{-ik_{i+1}a_i} \end{pmatrix} & \left(K(V_{i+1})\right) &= \begin{pmatrix} 1 & 1 \\ ik_{i+1} & -ik_{i+1} \end{pmatrix}
\end{aligned}
\tag{7}
$$

In Fig. 3, the four matrices denoted collectively as M_i account for transit through the region of potential V_i. It turns out that the expression for M_i assumes an especially simple form. Using eqn (7) to express M_i gives

$$
M_i = \begin{pmatrix} 1 & 1 \\ ik_i & -ik_i \end{pmatrix}\begin{pmatrix} e^{ik_i a_{i-1}} & 0 \\ 0 & e^{-ik_i a_{i-1}} \end{pmatrix}\begin{pmatrix} e^{-ik_i a_i} & 0 \\ 0 & e^{ik_i a_i} \end{pmatrix}\frac{1}{2}\begin{pmatrix} 1 & -i/k_i \\ 1 & i/k_i \end{pmatrix}. \tag{8}
$$

Defining δ_i as the distance over which V_i is constant (for example, $\delta_i = a_i - a_{i-1}$ in Fig. 3), and multiplying out the matrices in eqn (8) yields, following some algebra,

$$
M_i = \begin{pmatrix} \cos k_i\delta_i & -k_i^{-1}\sin k_i\delta_i \\ k_i \sin k_i\delta_i & \cos k_i\delta_i \end{pmatrix} \qquad E > V_i\ . \tag{9}
$$

As mentioned earlier, if a region is encountered in which the energy E is smaller than the potential, the substitution $k_i \to i\beta_i$ gives the appropriate matrix. For example, putting this substitution into eqn (9) yields

$$
M_i = \begin{pmatrix} \cosh \beta_i\delta_i & -\beta_i^{-1}\sinh \beta_i\delta_i \\ -\beta_i \sinh \beta_i\delta_i & \cosh \beta_i\delta_i \end{pmatrix} \qquad E < V_i\ . \tag{10}
$$

With the above tools, a series of steps is handled by writing M-matrices appropriate to each segment [see Fig. 3 and Fig. 4(a)]:

$$
\begin{pmatrix} A_L \\ B_L \end{pmatrix} = \left(E(V_L;a_0)\right)^{-1}\left(K(V_L)\right)^{-1} M_1 M_2 \ldots M_n \left(K(V_R)\right)\left(E(V_R;a_n)\right)\begin{pmatrix} A_R \\ B_R \end{pmatrix} \tag{11}
$$

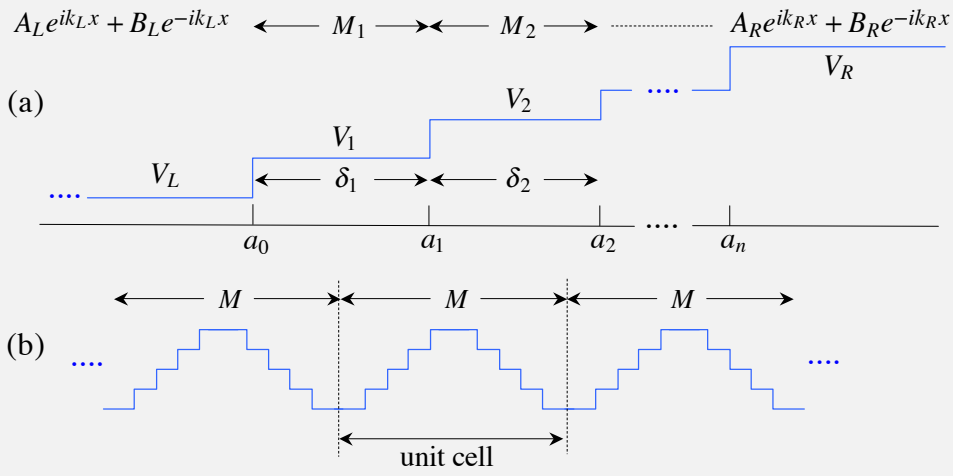


Figure 4. (a) Each flat region (except at the beginning and end) is characterized by its M matrix. (b) For periodic potentials, one need only deal with a single unit cell.

When the potential is periodic, say N sites on a ring [Fig. 4(b)], eqn (11) yields a neat form. Because the system closes on itself, $A_L = A_R$ and $B_L = B_R$. Referring to eqn (11), the two matrices at the end annihilate the two in the beginning. Make sure you understand how this works. The resulting M matrix need only be evaluated for one segment (unit cell), and we have

$$M^N = 1, \tag{12}$$

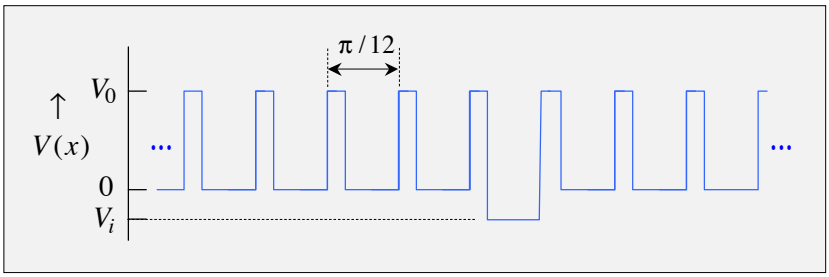
where N is the number of identical sites, each of which is described by the matrix M. Even though M for a repeating segment might involve many M_i, the mathematical problem does not depend on N. Thus, only a simple eigenvalue problem for a single segment needs to be solved. The N roots of eqn (12) yield the site-to-site phase differences that characterize the N solutions.

This completes the derivation. The method is easily implemented. For classic examples such as rectangular barrier, rectangular well, and single step, this method offers little or no advantage over algebraic derivations. For more interesting shapes, especially with periodic potentials, it is the method-of-choice. In addition, it can be used with a number of models such as tunneling through barriers, impurities in lattices (below), and Bragg scattering.

Impurities

It is interesting to examine what happens when a site is altered. This is a toy model for the substitution of an impurity in an otherwise periodic lattice. The number of sites is taken to be 24 – large enough to illustrate incipient features of the infinite lattice. The sites are identical except one, which differs only in binding strength. When V_i is negative

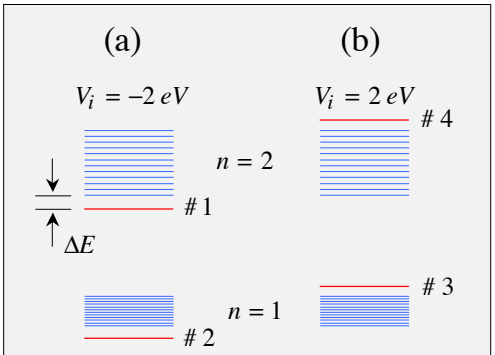
(as shown in the sketch below) the site is more binding. When V_i is positive the site is less binding.



What can be expected? A more-strongly-binding site adds a level below the band that arises from the 23 identical sites. Likewise, a less-strongly-binding site adds a level above the band from the 23 sites. The sketch below is for $m = m_e$, $V_0 = 10$ eV, 1 Å barrier, 4 Å well, and V_i values: + 2 eV and – 2 eV.

So far we have dealt with single-particle energies, and it has been assumed that electrons can enter the states without changing qualitatively their nature. Depending on the number of electrons, different things can happen. For example, a semiconductor might have the $n = 1$ (valence) band filled and the $n = 2$ (conduction) band nearly empty.

Referring to the sketch on the right, if the $n = 1$ band is filled and the $n = 2$ band is empty, an electron in state #1 is easily promoted to the $n = 2$ band as long as ΔE is not much larger than kT. It requires some energy to go to the $n = 2$ band, but entropy wins, and the electron spends most of its time in the $n = 2$ band. In this case, the impurity acts as a donor. It gives its electron to the band above it. Semiconductors that have impurities that donate electrons to a higher energy empty (conduction) band are n-type. In (b) an impurity level just above a filled (valence) band with no electron will acquire an electron from the filled band. This creates a hole in the otherwise filled band. This kind of impurity is p-type.



The 23 identical sites lead to bands (blue). The impurity results in levels (red) that lie either above or below the blue bands, depending on the nature of the impurity

In examining electron wave functions, it is found that the $n = 1$ and $n = 2$ states extend throughout the lattice. On the other hand, electron wave functions for red states are localized in the vicinities of their binding sites. State #1 can be like an exciton. In 3D lattices an electron bound to the vicinity of its positive charge (hole) is a Frenkel exciton. It migrates by hopping between impurities. If the electron is delocalized over many sites it is a Wannier-Mott exciton. It moves as a wave.

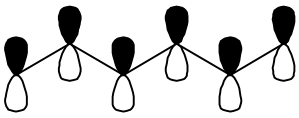
Chapter 3.

Hückel Model

Horse in Fog
Robert Wittig

Only in the U. S. will you find people who think the moon landing was fake and wrestling is real.

M. Cartmill



Contents

Preliminary Comments

In this chapter, atom-centered electron orbitals and their nearest neighbor interactions are discussed. The atoms are evenly spaced and, for the most part, arranged in closed or open chains. A great deal of correspondence will be revealed between the results obtained in this chapter and those obtained in Chapters 1 and 2. This chapter is designed for easy reading – front to back in one sitting.

In the 1930's, Erich Hückel developed a model that provides qualitative understanding of conjugated, aromatic, and anti-aromatic π-electron systems. It describes delocalized π-electron systems in terms of atomic *p*-orbitals centered on each of the participating atoms. These atoms are taken here (and in most cases) to be carbon. The atomic *p*-orbitals are oriented perpendicular to the plane that contains the carbon atoms, and it is assumed that their bonding can be treated independently of other orbitals such as the σ-bonded network that serves as the platform for the π-bonding. The π-system exists in the field of the substrate, and many interactions are implicit, entering through parameters.

This model is exceptional in its willingness to dispatch of nuance, and even important physics, to unearth effects. The Hamiltonian matrix is expressed in terms of just two parameters: one (α) for diagonal elements, the other (β) for nonzero off-diagonal elements. A third parameter is often introduced to account for overlap of adjacent atomic orbitals. Not much would be gained in the present treatment, so we will stick with just the two parameters, α and β.

What has come to be called Hückel theory is amazing. Though primitive and naïve, estimates can be made for complex systems in minutes, and the qualitative trends it predicts are usually correct. The model is empirical. It is sometimes referred to as semi-empirical, but parameters like α and β are obtained from experiment, so I prefer to think of it as empirical. Extensions such as the one introduced by Parr, Pariser, and Pople attest to the virtues of the original work, and extensive research by Hoffmann and coworkers using Hückel type models has impacted our understanding of solid state chemistry.

Nowadays, a high level of molecular orbital theory is available. However, the Hückel model will be used because it is transparent, easy to implement, and features that arise due to symmetry do not change with better theory. In so limited a venue as one modest chapter in these notes we are not motivated to enter the realm of electronic structure theory (Krylov group).

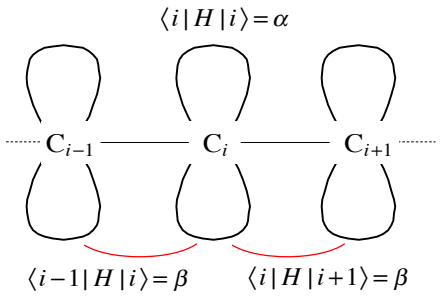
A cursory introduction is presented below. Following this, the model is applied to small hydrocarbons: allyl, cyclopropenyl, butadiene, cyclobutadiene, and benzene. Next, general expressions are derived for the energies and orbitals of open and closed chains comprised of even and odd numbers of carbon atoms. Neat pictorial ways to display the results are given and the large-chain limit is established. Our considerations are limited to

conjugated chains that are open or closed at the ends. We shall forego 2D networks except in the exercises and projects.

An uncanny congruence between the Hückel and POR models will be revealed. For example, we will again see that undulating site-to-site probability can be explained in an intuitive way. Likewise, the limit of a large number of atoms will be seen to give the conduction band of the metallic state, and so on.

The main idea is depicted in Fig. 1, where atomic sites are assigned indices i. In the atomic orbital basis, χ_i, each of the diagonal matrix elements, $\langle \chi_i \,|\, H \,|\, \chi_i \rangle$, is set equal to a constant, α, which is assumed here to be negative, because in zero order the electron is bound to the carbon atom.

Figure 1. The π-electron system of a hydrocarbon chain (open or closed) can be modeled using the Hückel Hamiltonian with a p-orbital on each atom. There are two parameters: the diagonal matrix element α, and the nearest-neighbor off-diagonal matrix element β, each of which is negative.



Environmental effects are ignored in the sense that the same α value is assigned to each atom that contributes a p-electron to the π-electron system, for example, regardless of whether the atom lies at the end or in the middle of an open chain. Each nearest neighbor interaction is assigned a nonzero matrix element whose value is β, which is also assumed to be negative. Matrix elements for more distant interactions such as $\langle i \,|\, H \,|\, i \pm 2 \rangle$, $\langle i \,|\, H \,|\, i \pm 3 \rangle$, and so on are set equal to zero. It is assumed tacitly that the system is planar, or effectively so, in which case it is assumed that the p-orbitals are displaced horizontally but not vertically from one another, nor are they tilted with respect to one another.

3.1. The Model

The essence of the model is given above. Let us now express this mathematically. In the atomic orbital basis, the characteristic (secular) equation is

$$\begin{pmatrix} H_{11}-ES_{11} & H_{12}-ES_{12} & H_{13}-ES_{13} & \cdots & H_{1N}-ES_{1N} \\ H_{21}-ES_{21} & H_{22}-ES_{22} & H_{23}-ES_{23} & \cdots & H_{2N}-ES_{2N} \\ H_{31}-ES_{31} & H_{32}-ES_{32} & H_{33}-ES_{33} & \cdots & H_{3N}-ES_{3N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ H_{N1}-ES_{N1} & H_{N2}-ES_{N2} & H_{N3}-ES_{N3} & \cdots & H_{NN}-ES_{NN} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_N \end{pmatrix} = 0. \qquad (3.1)$$

When the parameters α and β are introduced, this becomes

$$\begin{pmatrix} \alpha - ES_{11} & \beta - ES_{12} & -ES_{13} & \cdots & \beta\delta_{\text{cyc}} - ES_{1N} \\ \beta - ES_{21} & \alpha - ES_{22} & \beta - ES_{23} & \cdots & -ES_{2N} \\ -ES_{31} & \beta - ES_{32} & \alpha - ES_{33} & \cdots & -ES_{3N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \beta\delta_{\text{cyc}} - ES_{N1} & -ES_{N2} & -ES_{N3} & \cdots & \alpha - ES_{NN} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_N \end{pmatrix} = 0 \qquad (3.2)$$

where δ_{cyc} is 1 for closed chains and 0 for open chains. This parameter accounts for the fact that in a closed chain each atom has two nearest neighbors, whereas in an open chain the terminal atoms do not interact with one another.

The S_{ij} are overlap integrals for the atomic orbitals. For example, $S_{13} = \langle \chi_1 | \chi_3 \rangle$ is the overlap integral for orbitals centered at the first and third atoms. Referring to eqn (3.2), it is assumed that the terms $-ES_{ij}$ are equal to zero for $i \neq j$, and that they are equal to $-E$ for $i = j$. For normalized atomic orbitals, $S_{ii} = 1$. Because the atomic orbitals are taken as localized at their respective atoms, it is reasonable to assume that S_{ij} is zero for other than nearest neighbors, for example, $S_{13} = S_{14} = S_{15} \ldots = 0$ (see Fig. 2). But what about nearest neighbor terms like S_{12}: it is obvious that $S_{12} \neq 0$, otherwise there would be no bonding. Yet the model discards S_{12}.

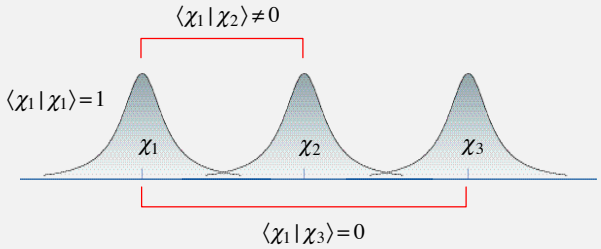


Figure 2. Electron orbitals χ_1, χ_2, and χ_3 are centered on C_1, C_2, and C_3, respectively. The overlap integral $\langle \chi_1 | \chi_2 \rangle$ is nonzero. The assumption $\langle \chi_1 | \chi_3 \rangle = 0$ is of good accuracy for the orbitals shown.

This raises an interesting point. The term H_{12} that has been assigned the value β in fact contains S_{12} because one part of $\langle \chi_1 | H | \chi_2 \rangle$ is $\alpha \langle \chi_1 | \chi_2 \rangle = \alpha S_{12}$. When S_{12} is discarded even though it is nonzero, its effect is accounted for, to the extent possible, by the parameter β. As mentioned earlier, better accuracy is achieved if overlap is taken explicitly into account with an additional parameter. Without this additional parameter, all nearest neighbor overlap effects are subsumed into β.

Using $S_{ij} = 0$ for $i \neq j$ in eqn (3.2) yields

$$\begin{pmatrix} \alpha - E & \beta & 0 & \cdots & \beta\delta_{\text{cyc}} \\ \beta & \alpha - E & \beta & \cdots & 0 \\ 0 & \beta & \alpha - E & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \beta\delta_{\text{cyc}} & 0 & 0 & \cdots & \alpha - E \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_N \end{pmatrix} = 0. \qquad (3.3)$$

Equation (3.3) is easily solved using software such as *Mathcad* and *Mathematica*. Though a useful closed form solution is derived below, planar networks and variants in which overlap is taken into explicit account do not, in general, have closed form solutions. Thus, you might as well get used to packages such as Mathcad and Mathematica.

Allyl and Cyclopropenyl

In allyl (C_3H_5), three electrons participate in π-bonding. The characteristic equation is

$$\begin{pmatrix} \alpha - E & \beta & 0 \\ \beta & \alpha - E & \beta \\ 0 & \beta & \alpha - E \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = 0 \,. \tag{3.4}$$

Setting the determinant to zero: $(\alpha - E)^3 - 2(\alpha - E)\beta^2 = 0$, yields eigenvalues: $\alpha + \sqrt{2}\beta$, α, and $\alpha - \sqrt{2}\beta$, and the corresponding orbitals are

$$\alpha + \sqrt{2}\beta : \frac{1}{2}\begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix} \qquad \alpha : \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \qquad \alpha - \sqrt{2}\beta : \frac{1}{2}\begin{pmatrix} -1 \\ \sqrt{2} \\ -1 \end{pmatrix} . \tag{3.5}$$

Figure 3 indicates these orbitals. From bottom to top (increasing energy), they are bonding, non-bonding, and antibonding. When three electrons are put into the orbitals that constitute the ground state, two go into the bonding orbital and one goes into the non-bonding orbital. This non-bonding orbital looks like perhaps it should be called anti-bonding. However, the distance between the terminal atoms is large, so there is little interaction between these orbitals and they are therefore called non-bonding.

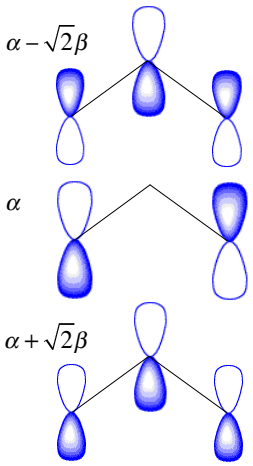


Figure 3. Allyl: bottom to top, bonding, non-bonding, anti-bonding.

Thus, the ground state has a sufficiently low energy that allyl is reasonably stable. It serves as an intermediate in a number of reactive environments.

Whereas the terminal atoms in allyl are not bonded directly to one another, in cyclopropenyl (C_3H_3) each atom is bonded to two nearest neighbors. Thus, competing effects are anticipated. The delocalization provided by the closed chain lowers the energy of the bonding orbital relative to allyl. However, considerable strain energy is needed to create the σ-bonded equilateral triangle. The presence of matrix elements β in the upper right and lower left corners of the Hamiltonian matrix is the only mathematical difference between open and closed chains. It is interesting that the same form for the characteristic equation applies to such different cyclic systems as C_3H_3, Na_3, and N_3.

The characteristic equation for cyclopropenyl is

$$\begin{pmatrix} \alpha - E & \beta & \beta \\ \beta & \alpha - E & \beta \\ \beta & \beta & \alpha - E \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = 0 \,. \tag{3.6}$$

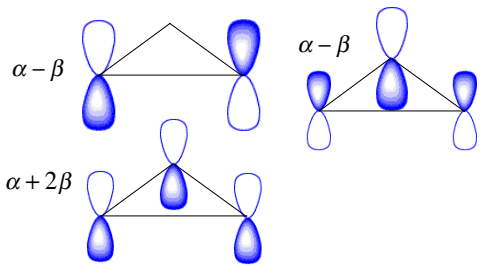
The eigenvalues are $\alpha + 2\beta$ and a degenerate pair: $\alpha - \beta$. Corresponding orbitals are

$$\alpha + 2\beta : \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \qquad \alpha - \beta : \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \qquad \alpha - \beta : \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \tag{3.7}$$

The degenerate (*e*-type) orbitals are not unique, as any linear combination of them is a solution of the Schrödinger equation. This is a general result for closed chains. Namely, energies appear in degenerate pairs except for the lowest and (for even numbers of atoms) the highest energy orbitals. This is due to the clockwise and counterclockwise phase progressions around the closed paths. Each must have the same energy.

The cyclopropenyl orbitals in Fig. 4 have things in common with those of allyl. In each, the ground state consists of two electrons in the bonding orbital and one in the next highest orbital. In allyl, the third electron goes into the non-bonding orbital. However, in cyclopropenyl the two orbitals at the significantly higher energy of $\alpha - \beta$ are degenerate, which makes this case less straightforward. It is well known that the Hückel model fails to describe cyclopropenyl. This is not a trivial problem even at a higher level of theory. For example, the fact that the D_{3h} ground state is degenerate enables the equilateral triangle to distort to C_{2v} via the Jahn-Teller effect. The Hückel model does not treat the Jahn-Teller effect, but it alerts us to the fact that a higher level of theory is required.

Figure 4. Cyclopropenyl (C_3H_3) π-orbitals: the lowest orbital is much more stable than the other two. The orbitals at energy $\alpha - \beta$ are degenerate. Note that the cation $C_3H_3^+$ (cyclopropenium) is relatively stable because both electrons are in the bonding orbital.



Though the π-system energy of cyclopropenyl is lower than that of allyl ($3\alpha + 3\beta$ versus $3\alpha + 2\sqrt{2}\beta$, respectively), this does not mean the former is more stable. The strain energy needed to form the triangle raises considerably the energy of the σ-bonds. Consequently, cyclopropenyl is unstable, in part because of its high energy degenerate HOMO. On the other hand, cyclopropenium ($C_3H_3^+$) has two bonding electrons, so it is relatively stable.

Butadiene and Cyclobutadiene

The characteristic equation for butadiene and cyclobutadiene is given in eqn (3.8), and the energies and orbitals are given in the box, eqns (3.9) and (3.10).

$$\begin{pmatrix} \alpha - E & \beta & 0 & \beta\delta_{\text{cyc}} \\ \beta & \alpha - E & \beta & 0 \\ 0 & \beta & \alpha - E & \beta \\ \beta\delta_{\text{cyc}} & 0 & \beta & \alpha - E \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} = 0 \tag{3.8}$$

Equations (3.9) and (3.10) predict that the π-system of butadiene is lower in energy than that of cyclobutadiene by 0.47β. Together with the strain energy engendered by the four-membered ring, this is consistent with the fact that square cyclobutadiene is unstable. Its degenerate orbitals ($E_2 = E_3 = \alpha$) enable the ground state to distort spontaneously into a rectangular shape via the Jahn-Teller effect. This has been verified experimentally.

Butadiene Orbitals

$$E_1 = \alpha + 1.618\beta \quad E_2 = \alpha + 0.618\beta \quad E_3 = \alpha - 0.618\beta \quad E_4 = \alpha - 1.618\beta$$

$$\phi_1 = \begin{pmatrix} 0.372 \\ 0.602 \\ 0.602 \\ 0.372 \end{pmatrix} \quad \phi_2 = \begin{pmatrix} 0.602 \\ 0.372 \\ -0.372 \\ -0.602 \end{pmatrix} \quad \phi_3 = \begin{pmatrix} 0.602 \\ -0.372 \\ -0.372 \\ 0.602 \end{pmatrix} \quad \phi_4 = \begin{pmatrix} -0.372 \\ 0.602 \\ -0.602 \\ 0.372 \end{pmatrix} \tag{3.9}$$

Cyclobutadiene Orbitals

$$E_1 = \alpha + 2\beta \quad E_2 = E_3 = \alpha \quad E_2 = E_3 = \alpha \quad E_4 = \alpha - 2\beta$$

$$\phi_1 = \begin{pmatrix} 0.500 \\ 0.500 \\ 0.500 \\ 0.500 \end{pmatrix} \quad \phi_2 = \begin{pmatrix} 0 \\ -0.707 \\ 0 \\ 0.707 \end{pmatrix} \quad \phi_3 = \begin{pmatrix} 0.707 \\ 0 \\ -0.707 \\ 0 \end{pmatrix} \quad \phi_4 = \begin{pmatrix} 0.500 \\ -0.500 \\ 0.500 \\ -0.500 \end{pmatrix} \tag{3.10}$$

Benzene

Benzene is the quintessential example of the stabilization of a π-system through electron delocalization. Its characteristic equation is

$$\begin{pmatrix} \alpha-E & \beta & 0 & 0 & 0 & \beta \\ \beta & \alpha-E & \beta & 0 & 0 & 0 \\ 0 & \beta & \alpha-E & \beta & 0 & 0 \\ 0 & 0 & \beta & \alpha-E & \beta & 0 \\ 0 & 0 & 0 & \beta & \alpha-E & \beta \\ \beta & 0 & 0 & 0 & \beta & \alpha-E \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \end{pmatrix} = 0 . \tag{3.11}$$

The eigenvalues, in order of increasing energy, are: $\alpha+2\beta$, $\alpha+\beta$ (doubly degenerate), $\alpha-\beta$ (doubly degenerate), and $\alpha-2\beta$. As mentioned many times, the wave functions for the degenerate orbitals are not unique. Given a pair of orbital wave functions for a doubly degenerate case, any linear combination of them is a solution. The orbitals for the degenerate pairs that you find most often in books are the real functions:

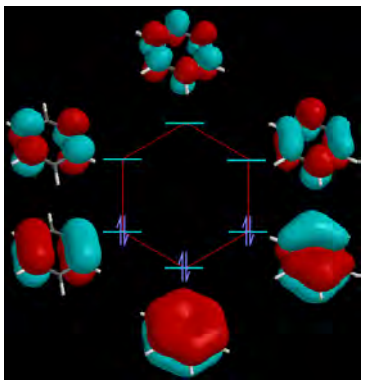
$$\alpha+\beta: \quad \phi_2 = \chi_2 + \chi_3 - \chi_5 - \chi_6 \tag{3.12}$$

$$\phi_3 = 2\chi_1 + \chi_2 - \chi_3 - 2\chi_4 - \chi_5 + \chi_6 \tag{3.13}$$

$$\alpha-\beta: \quad \phi_4 = \chi_2 - \chi_3 + \chi_5 - \chi_6 \tag{3.14}$$

$$\phi_5 = 2\chi_1 - \chi_2 - \chi_3 + 2\chi_4 - \chi_5 - \chi_6 \tag{3.15}$$

The molecular orbitals are given without normalization factors, as the latter need to take atomic overlap into account.



The picture on the right shows the six real molecular orbitals. The degenerate ones are described by the above equations. In the ground state, the electrons occupy the $\alpha+2\beta$ orbital (lowest energy, 2 electrons) and the $\alpha+\beta$ orbital (doubly degenerate, 4 electrons).

The cases, treated so far of 3, 4, and 6 sites, with their energy bounds of $\pm 2\beta$ and trigonometric-looking numbers for expansion coefficients, invite ponder: Might a pattern be emerging? How do results vary with the number of atoms? How shall Hückel theory be implemented in the limit of an infinite lattice?

It is time to derive a few general results.

3.2. Closed chains

For a closed chain the characteristic equation (3.3) can be written

$$c_{n-1} + \left(\frac{\alpha - E}{\beta} \right) c_n + c_{n+1} = 0 \, . \tag{3.16}$$

Because the *p*-orbitals are taken to be equivalent, it must be possible to express the molecular orbitals in terms of *p*-orbitals that differ from site-to-site only by a phase. Thus, leaving aside normalization, each expansion coefficient is of the form

$$c_n = e^{i2\pi kn/N} \, . \tag{3.17}$$

In the argument of the exponential, N is the number of atoms, n is the location of a given atom, and k denotes the k^{th} molecular orbital. Because the number of *p*-orbitals is N, there are N molecular orbitals. Thus, k has a total of N values. The only way the different molecular orbitals can be manifest in the coefficients is in phase progression.

The k values in eqn (3.17) must be integers. For example, for $N = 6$, k has values 0, ± 1, ± 2, and 3. Notice that the 3 is not accompanied by $\pm$. A combination of $+3$ and -3 can be used as long as we end up with a single molecular orbital (recall the black points in Fig. 7 of Chapter 1). Putting eqn (3.17) into eqn (3.16) yields

$$e^{i2\pi k(n-1)/N} + \left(\frac{\alpha - E}{\beta} \right) e^{i2\pi kn/N} + e^{i2\pi k(n+1)/N} = 0 \, . \tag{3.18}$$

Dividing by $e^{i2\pi kn/N}$ gives

$$e^{-i2\pi k/N} + e^{i2\pi k/N} = \left(\frac{E - \alpha}{\beta} \right) . \tag{3.19}$$

Thus, the energy of the k^{th} molecular orbital is

Closed Chains

$$E_k = \alpha + 2\beta \cos\left(\frac{2\pi}{N} k \right) \quad \text{where } k = 0, \pm 1, \pm 2 \ldots N/2. \tag{3.20}$$

The fact that energy is proportional to a cosine whose argument is 2π times a rational number invites a geometric construction. Figure 5 shows that the energies can be obtained by putting a regular polygon inside a circle whose radius is $2|\beta|$. The polygon is aligned such that one of its apices is at $\theta = 0$, where θ ($= 2\pi k / N$) is indicated in the upper left

entry. The energies are then read from the diagram. They are the points of contact of each of the apices. Double degeneracy is present automatically.

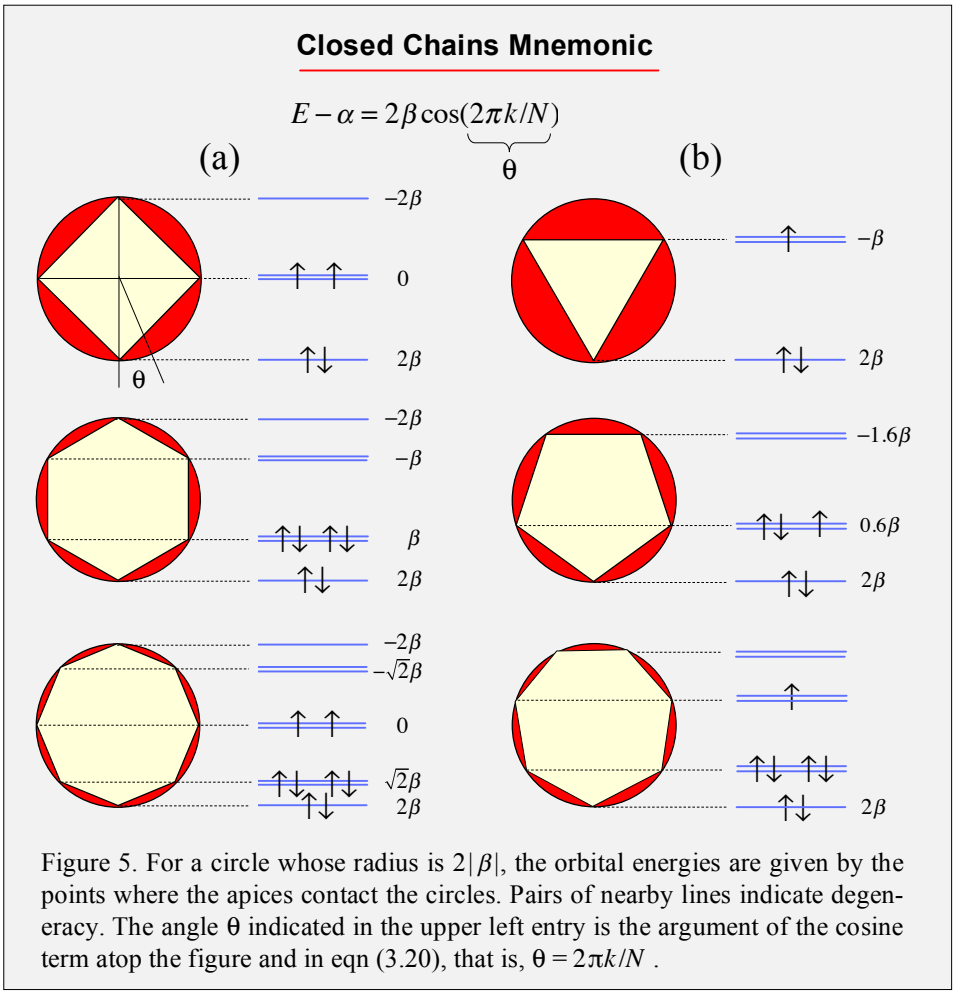


Figure 5. For a circle whose radius is $2|\beta|$, the orbital energies are given by the points where the apices contact the circles. Pairs of nearby lines indicate degeneracy. The angle θ indicated in the upper left entry is the argument of the cosine term atop the figure and in eqn (3.20), that is, $\theta = 2\pi k/N$.

Figure 5(a) (left column) is for even N. From top to bottom, the molecules are cyclobutadiene, benzene, and cyclooctatetraene. These diagrams illustrate the quandary faced by cyclobutadiene and cyclooctatetraene of having to put electrons in degenerate nonbonding orbitals, notably, versus the considerable stability of benzene. Furthermore, we see that stable situations occur with the successive addition of 4 atoms: 6, 10, 14, 18…. Thus, a trend is identified in which conjugated closed chains consisting of $4n + 2$ atoms are stable. These are called aromatic. Those with $4n$ atoms are less stable. These are called anti-aromatic. For example, cyclobutadiene and cyclooctatetraene distort to rectangular and boat shapes, respectively. These propensities are summarized in Hückel's rule:

Hückel's Rule

Cyclic planar molecules in which each atom has a p-orbital are aromatic if they contain $4n+2$ π-electrons.

Figure 5(b) (right column) is for odd N: from the top, cyclopropenyl (C_3H_3), cyclopentadienyl (C_5H_5), and cycloheptatrienyl (C_7H_7). These are radicals, as they have unpaired electrons. The instability of cyclopropenyl comes as no surprise in light of the high HOMO energy and strain of the σ-bonded frame. Note that, in general, radicals have more low-energy spectroscopic transitions than stable molecules because electrons can be promoted from lower orbitals to the HOMO.

As N increases, differences between odd and even N become less important. Referring to Fig. 5, imagine a many-sided polygon embedded in a circle. Clearly, the distinction between results obtained with odd and even N is lost in the large-N limit. This was noted in Fig. 8 of Chapter 1. Eigenvalues for 12 and 13 sites are compared, and it is seen that there is no significant difference.

The geometric construction of Fig. 5 yields a conduction band in the large-N limit. Figure 6 illustrates this by showing a progression that begins with the N values used in Fig. 5. As N increases, though the number of degenerate pairs increases, energies are restricted to lie within the range $\pm 2\beta$. For $N = 24$, the energies are on their way to the formation of a continuous band. When N electrons are added (two per level), they fill the band up to $E = \alpha$. This half-filled band satisfies the criterion for a conductor.

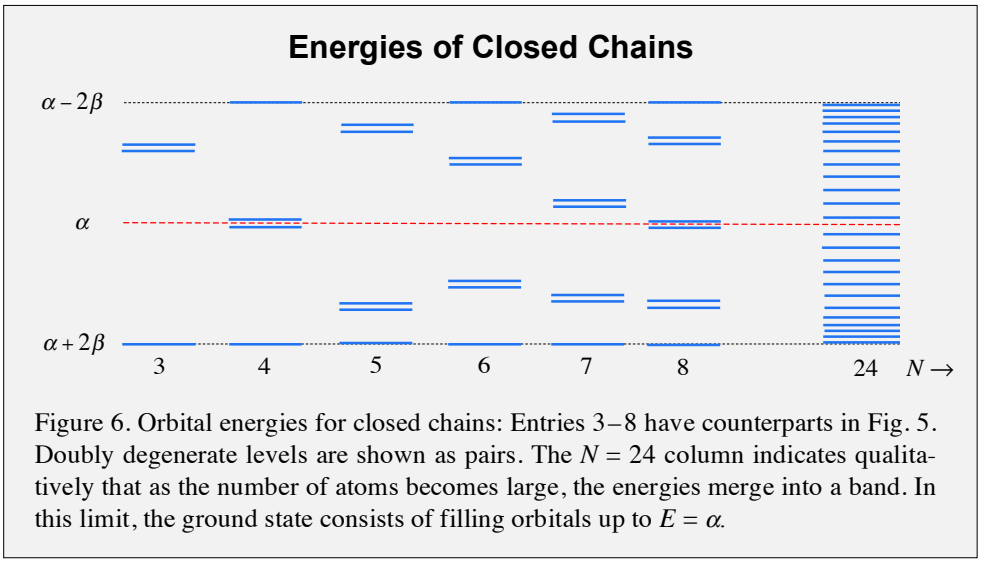


Figure 6. Orbital energies for closed chains: Entries 3–8 have counterparts in Fig. 5. Doubly degenerate levels are shown as pairs. The $N = 24$ column indicates qualitatively that as the number of atoms becomes large, the energies merge into a band. In this limit, the ground state consists of filling orbitals up to $E = \alpha$.

3.3. Open Chains

Now consider open chains. Just as the closed chain is analogous to the POR, the open chain is analogous to the particle-in-a-box (PIB) with infinite potential at the walls. The strategy for obtaining the orbital energies and wave functions is not much different than in the case of closed chains. The difference arises from the boundary conditions. The POR requires $\psi(2\pi) = \psi(0)$, whereas the PIB requires that the wave function vanishes at the walls.

For open chains, the orbitals must be standing waves that vanish at the edges of the chain. But what are these edges? The terminal atoms cannot serve as edges, or the orbitals would be zero at these locations. This is impossible. The charge densities of the molecular orbitals must add up such that for the ground state of the system each carbon atom binds one electron. Consequently, these so-called edges must lie outside the real atoms, as indicated in Fig. 7.

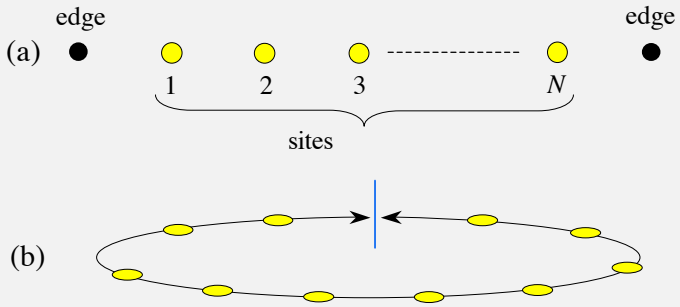


Figure 7. (a) An open chain of N atoms: To satisfy the boundary conditions, edges (black dots) are placed on each side of the chain. (b) The same effect is achieved with the particle on a ring; an impenetrable membrane (blue) serves as the edges.

Phase progresses uniformly from one site to the next, analogous to the closed chain, except in the present case the coefficients are expressed in terms of standing waves that vanish at the edges. The number of intervals between the edges (Fig. 7) is $N + 1$, and the expressions analogous to eqns (3.17) and (3.18) are

$$c_n = \sin\left(\frac{\pi kn}{N+1}\right) \qquad k = 1, 2, 3 \ldots N \tag{3.21}$$

and

$$\sin\left(\frac{\pi k(n-1)}{N+1}\right) + \left(\frac{\alpha - E}{\beta}\right)\sin\left(\frac{\pi kn}{N+1}\right) + \sin\left(\frac{\pi k(n+1)}{N+1}\right) = 0 \tag{3.22}$$

Applying the trigonometric identity: $\sin(a+b) = \sin a \cos b + \cos a \sin b$, to the first and third terms, and dividing everything by $\sin\left(\pi kn/(N+1)\right)$ gives

Open Chains

$$E_k = \alpha + 2\beta\cos\left(\frac{\pi k}{N+1}\right) \qquad k = 1, 2 \ldots N \tag{3.23}$$

This has the same form as eqn (3.20), except for the argument of the cosine and the range of k values. The energies are non-degenerate, spanning the range ± 2β. Figure 8 is analogous to Fig. 6. It gives results for open chains of 2, 3, 4, 5, 6, 12, and 24 atoms,

showing how the system approaches the large-N limit in which the discrete levels merge into a band.

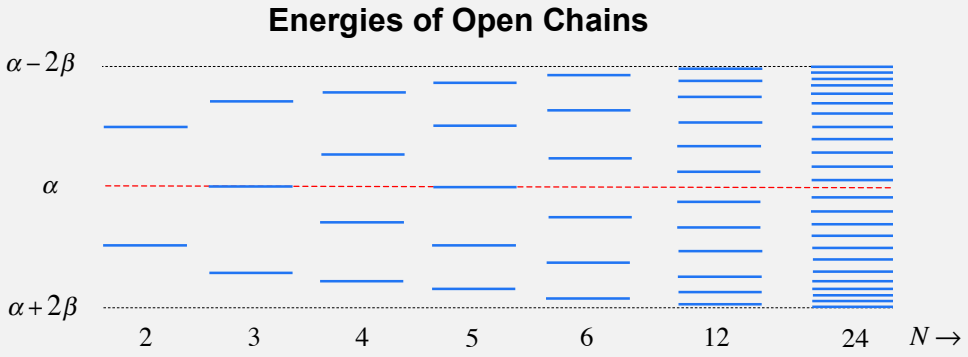


Figure 8. Orbital energies for open chains: For small N, open and closed chains differ qualitatively (compare to Fig. 6), while for large N there is no significant difference.

In the case of the open chain, though we cannot use the same mnemonic for the energies as the one given in Fig. 5, it is easy to devise a similar one. Figure 9 is based on eqn (3.23) in the same way that Fig. 5 is based on eqn (3.20). Figure 9 is merely a pictorial realization of eqn (3.23). To accommodate N atoms requires $N + 2$ points. The two at the ends (the edges) do not enter the expression for E_k given by eqn (3.23). The others are simple projections as shown in the figure. Note that the levels are non-degenerate.

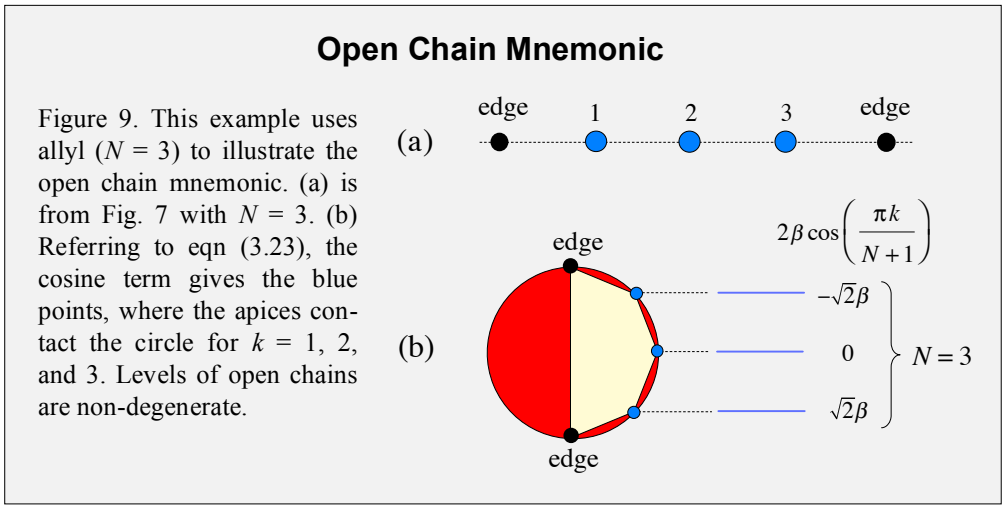


Figure 9. This example uses allyl (N = 3) to illustrate the open chain mnemonic. (a) is from Fig. 7 with N = 3. (b) Referring to eqn (3.23), the cosine term gives the blue points, where the apices contact the circle for k = 1, 2, and 3. Levels of open chains are non-degenerate.

Figures 5 and 9 summarize the conjugated chain energies: odd/even; open/closed; long/short. Equation (3.17) ($c_n = e^{i2\pi kn/N}$) gives the expansion coefficients for molecular orbitals for closed chains. The degenerate pair orbitals can also be expressed as real functions, as was done in eqns (3.12)−(3.15). For open chains, the molecular orbitals in terms of atomic constituents are given by eqn (3.21): $c_n = \sin[\pi kn/(N+1)]$. The values can be read from the pictures in Fig. 10.

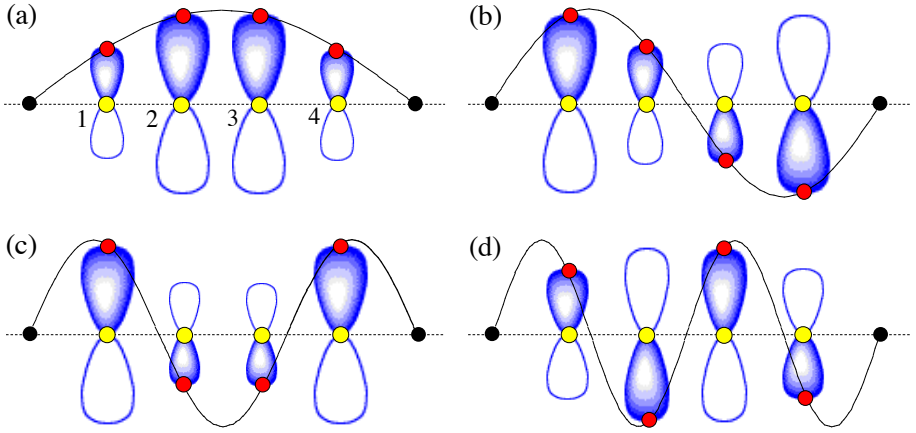
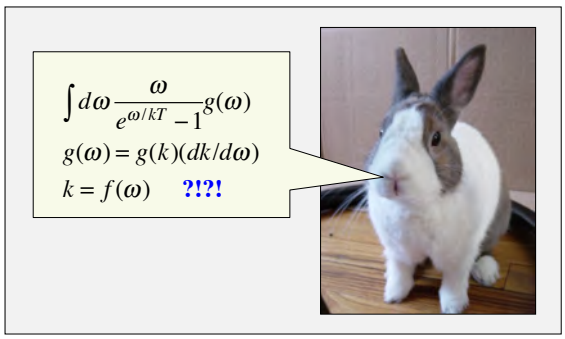


Figure 10. Open chain of 4 atoms: The equation $c_n = \sin\left(\pi kn/(N+1)\right)$ displays the molecular orbitals in terms of the atomic orbitals. Entries (a)−(d) are for k = 1−4. Red points indicate values of c_n at respective locations n.

3.4. Density of States

If this is your first encounter with density of states, you are advised to treat it with respect. It is not sophisticated, nor does it require complicated math. However, in the beginning it can be subtle. We will start with the familiar PIB and POR examples. The densities of states will then be calculated for the open and closed chains of the Hückel model. We will end with densities in wave number space (*k*-space).



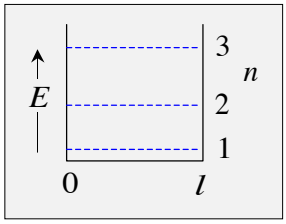
Particle-in-a-Box

The first example is a particle in a 1D box having impenetrable walls. The expressions for the eigenvalues and derivative with respect to quantum number n are

$$E = \frac{h^2}{8ml^2} n^2 \tag{3.24}$$

$$\frac{dE}{dn} = \frac{h^2}{4ml^2} n \tag{3.25}$$

As used here, the term density of states means the number of states per unit energy: dn/dE. Because the PIB states are not degenerate, the number of n values in the range ΔE is identical to the number of states in this range. Referring to eqn (3.25), we shall treat n as a continuous variable. This is acceptable, but be sure to keep in mind the discrete nature of the levels.



The density of energy states is not the only density we will encounter, but it is a common one. For example, we might also have to deal with the number of states per unit momentum or per unit wave vector. However, in the present example, the word state is understood to mean energy state. The density of states we are considering, $\rho(E)$, is obtained directly from eqns (3.24) and (3.25):

$$\rho(E) = \frac{dn}{dE} = \frac{4ml^2}{h^2}\frac{1}{n} = \frac{4ml^2}{h^2}\left(\frac{h^2}{8ml^2E}\right)^{1/2}. \tag{3.26}$$

Thus,

$$\rho(E) = \frac{l}{h}(2m)^{1/2}E^{-1/2} \quad \text{particle-in-a-box.} \tag{3.27}$$

Because $\rho(E)$ is proportional to $E^{-1/2}$, it goes to infinity as $E \to 0$. Referring to Figs. 6 and 8, the columns for 24 atoms illustrate the fact that the density of states rises at the band edge. Leaving aside the mathematical infinity, it is obvious that nothing unphysical happens near the band edges. The mathematical infinity in fact is nothing more than a manifestation of the quadratic dependence of energy on n: level separations increase with n. For discrete n:

$$E_n - E_{n-1} = \frac{h^2}{8ml^2}(2n-1) \Rightarrow \frac{\Delta n}{\Delta E} = \frac{4ml^2}{h^2\left(n-\frac{1}{2}\right)},$$

whereas for continuous n:

$$\frac{dn}{dE} = \frac{4ml^2}{h^2 n}.$$

The singularity in eqn (3.27) is not worrisome. When $\rho(E)$ is integrated over an energy range, we obtain, by definition, the total number of states in that range. In Figs. 6 and 8, the number of levels in the band (excluding spin) is the number of atoms. For a PIB with infinite walls, we must choose a cutoff energy, because in nature there is no such thing as a potential with infinitely strong binding. In neither case, however, does the integral of $\rho(E)$ diverge. We have encountered what is called an integrable singularity.

Particle-on-a-Ring

Next, consider the particle-on-a-ring:

$$E = Bm^2 \tag{3.28}$$

$$\frac{dE}{dm} = 2Bm. \tag{3.29}$$

where B is the rotational constant. Again, using $\rho(E) = dm/dE$ gives

$$\rho(E) = \frac{1}{2B}\frac{1}{m} = \frac{1}{2B}\left(\frac{B}{E}\right)^{1/2} = \frac{1}{2}B^{-1/2}E^{-1/2}. \tag{3.30}$$

As in the case of the PIB, $\rho(E)$ is proportional to $E^{-1/2}$. However, something is clearly amiss! Equation (3.30) cannot be correct because the twofold POR degeneracy has not been taken into account. Except for m = 0, the m-states can be arranged in degenerate pairs. In other words, m values that have the same magnitude but differ in sign have the same energy. Therefore, it is necessary to multiply eqn (3.30) by 2. This yields

$$\rho(E) = B^{-1/2}E^{-1/2} \quad \text{particle-on-a-ring.} \tag{3.31}$$

To check the PIB and POR results for consistency, compare the coefficients of $E^{-1/2}$ in eqns (3.27) and (3.31), $(l/h)(2m)^{1/2}$ and $B^{-1/2}$, respectively. In doing this, you will see that these coefficients have the same value as long as the length of the box is equal to the circumference of the ring. A parabolic variation of E with quantum number always gives $E^{-1/2}$ dependence.

It is equally instructive to derive the densities of states for harmonic oscillators (1D, 2D…), bound states of the hydrogen atom, a particle on a sphere, symmetric and asymmetric tops, and other elementary examples. Let us now return to the Hückel model.

Hückel Density of States

Here we shall determine how the π system orbital energies are distributed throughout the band when the number of atoms is large. This distribution would be meaningless were the number of atoms small. For example, the HOMO-LUMO gap of benzene gives rise to an absorption deep in the ultraviolet. As N increases, the HOMO-LUMO gap decreases. In the large-N limit, the orbital energies merge into a band.

In the large-N limit, the density of states can be obtained using either eqn (3.27) or (3.31). These expressions are equivalent insofar as density of states is concerned, though twofold degeneracy must be taken into account when using eqn (3.20). Namely, k values that have the same magnitude but differ in sign have the same energy.

The quantum number k indexes the states, so to obtain the density of states it is necessary to obtain dk/dE. For non-degenerate states $\rho(E) = |dk/dE|$, while for degenerate states $|dk/dE|$ is multiplied by the degeneracy. The approach taken here is essentially the same as the one used with eqns (3.24) – (3.31). To proceed, let us differentiate eqn (3.23) and use the fact that $N+1$ can be replaced with N because N is large:

$$\frac{dE}{dk} = -\frac{2\pi\beta}{N}\sin\left(\frac{\pi k}{N}\right). \tag{3.32}$$

Minus signs are immaterial for density of states, which is always positive. Thus, the density of states is

$$\rho(E) = \frac{N}{2\pi|\beta|}\,\frac{1}{\sin(\pi k/N)}, \tag{3.33}$$

where $0 \le k \le N$.

The quantum number k is usually not as useful as the wave vector. Consequently, k will now be replaced with the wave vector, which we define as: $k' = 2\pi / \lambda$. From Fig. 10 we see that the quantum number k is the number of half wavelengths contained in the length L, that is, $k = 2L / \lambda$. To express the $\pi k / N$ argument of the sine in eqn (3.33) in terms of the wave vector k', use $k' = 2\pi / \lambda = (\pi / L)k$. Therefore, $\pi k / N = k'a$, and eqn (3.33) becomes

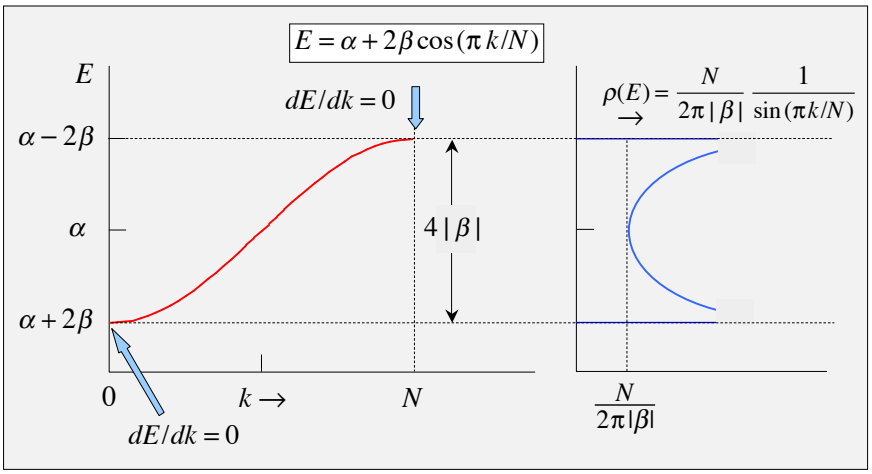
$$\rho(E) = \frac{N}{2\pi |\beta|} \frac{1}{\sin k'a}, \tag{3.34}$$

where: $0 \le k' \le \pi / a$.

The maximum value of k' is π / a, which corresponds to $k = N$. A qualitative sketch (educated guess) of $\rho(E)$ is given in Fig. 11. The upper and lower boxes are the same except that entries in the upper box are expressed in terms of quantum number k, whereas entries in the lower box are expressed in terms of wave vector k'. Note the singularities in the plots of $\rho(E)$ versus E that occur at $E = \alpha \pm 2\beta$. These are referred to as van Hove singularities.

You might wonder why E appears on the left hand side of eqn (3.33) in $\rho(E)$, whereas k appears on the right hand side, and likewise for eqn (3.34). After all, given that $\rho(E)$ is a function of E, should not the right hand side be a function of E rather than a function of k? On occasion (not with the Hückel model) $\rho(E)$ turns out to be independent of energy, in which case the right hand side is constant. This is fine, but again, why is the right hand side of eqn (3.33) expressed as a function of k?

The reason is that the k in eqn (3.33) is written with the understanding that it is to be expressed in terms of E. The relationship between k and E is given by eqn (3.23). This did not arise with the PIB and POR examples because the relationships between E and the respective quantum numbers were trivial. In other words, no notational convenience would have been achieved by retaining a quantum number on the right hand side of the $\rho(E)$ equations.



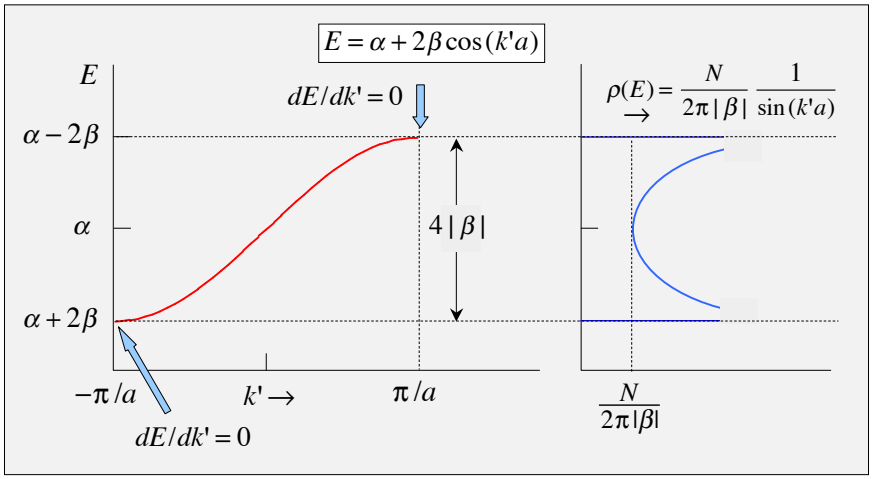


Figure 11. Upper left: E versus quantum number k. Upper right: sketch (qualitative) showing the density of states $\rho(E)$ given by eqn (3.33). The lower box is the same as the upper box except some entries are expressed in terms of wave vector k' instead of quantum number k. The van Hove singularities at $E = \alpha \pm 2\beta$ are indicated as open regions.

As mentioned earlier, the same result can be obtained using the closed chain. To see how this is done, first obtain dk/dE by differentiating eqn (3.20): $E = \alpha + 2\beta\cos(2\pi k/N)$, where: $-N/2 \le k \le N/2$. Figure 12 shows two equivalent regions in a plot of $|dk/dE|$ versus k. In other words, there is mirror symmetry about $k = 0$. To facilitate comparison with eqn (3.33), we can use just the $k > 0$ part of Fig. 12, as long as the number we obtain is multiplied by 2 when writing the density of states. This takes into account the twofold degeneracy. Thus,

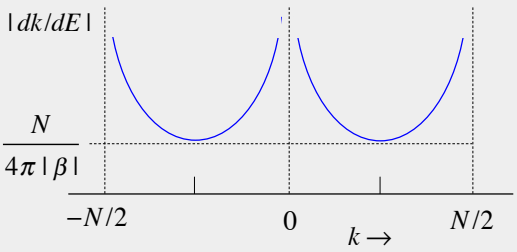


Figure 12. Two parts contribute equally to the density of states. We need to deal only with one part.

$$\rho(E) = \frac{N}{2\pi|\beta|}\,\frac{1}{\sin(2\pi k/N)}, \tag{3.35}$$

where: $0 \le k \le N/2$.

The argument of the sine ($2\pi k/N$) over the range $0 \le k \le N/2$, is equivalent to the argument of the sine in eqn (3.33), namely, $\pi k/N$, but over the larger range $0 \le k \le N$. Thus, eqns (3.35) and (3.33) are equivalent. What if, instead of the density of states in energy space, we want the density in *wave vector space*, namely, the number of states per unit wave vector. This is given by

$$\rho(k') = dk/dk'. \tag{3.36}$$

For the open Hückel chain, $\rho(k') = L/\pi$. From Fig. 10, we see that $k = 2L/\lambda = k'L/\pi$ $\Rightarrow$ $dk/dk' = L/\pi$. Note that this is the same as the PIB density of states. There, the boundary condition: $\sin k'x = 0$ at $x = L$, yields $k'L = n\pi$, in which case, $dn/dk' = L/\pi$. On the other hand, for the closed chain, $\rho(k') = L/2\pi$. Positive and negative k' are distinguished because they are distinct in wave vector space. The same result is, of course, obtained for the POR. Namely, the boundary condition $k'L = 2\pi k$ gives $\rho(k') = L/2\pi$.

It is straightforward to extend the 1D density in wave vector space to 2D and 3D. This is left as an exercise.

Chapter 4.

Infinite One-Dimensional Lattice

Hall of Windows
Robert Wittig

As an adolescent I aspired to lasting fame. I craved factual certainty, and I thirsted for a meaningful vision of human life – so I became a scientist. This is like becoming an archbishop so you can meet girls.

Contents

Comment

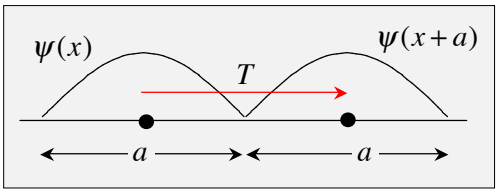
Periodic 1D lattices are examined here using a mathematical approach that is commonplace in solid-state physics, and that complements those used in Chapters 1 – 3. The number of sites can range from small to infinite. The math does not change. Though the mathematical approach of the present chapter differs from those used in Chapters 1 – 3, not surprisingly, similar results are obtained. My strategy is that the use of complementary perspectives facilitates the retention of concepts.

4.1. Bloch Waves

The first item on the agenda is the mathematical form that the eigenfunctions must assume as a consequence of the discrete lattice symmetry presented by evenly spaced sites. Toward this end, we start by introducing the translation operator, T, that brings about the transformation: $x \to x + a$, where a is the lattice spacing:

$$T\psi_k(x) = \psi_k(x+a). \tag{4.1}$$

Periodic boundary conditions shall be used throughout.



The operator T acts on a wave function that is defined over an interval of length a. It translates this wave function by the distance a, as indicated schematically on the right. The transformed wave function can acquire a phase but nothing more. Because T is a symmetry operation of the Hamiltonian, it commutes with H, and therefore T and H can have simultaneous eigenfunctions. The reason for labeling the eigenfunctions and eigenvalues of T and H with the index k will be discussed below. It is probably already clear, or at least not surprising, in light of the previous chapters. With the eigenvalues of T labeled λ_k we have

$$T\psi_k(x) = \lambda_k\psi_k(x). \tag{4.2}$$

The eigenvalues λ_k must satisfy $|\lambda_k| = 1$ to ensure that the probability density does not differ from site to site.[1] Thus, the λ_k can be expressed as

$$\lambda_k = e^{ika}. \tag{4.3}$$

[1] Consider $|\psi_k(x)|^2$ for the site interval: $0 \le x \le a$. We require that the probability density for another site is $|\psi_k(x-na)|^2$, where the integer n denotes the location of the other site.

In other words, $\psi_k(x)$ can only vary from site to site by a phase factor whose argument is proportional to a.

Equation (4.3) defines the wave vector k. Though k appears in combination with the lattice spacing a, the allowed values of k do not reflect the lattice periodicity, except in the indirect sense that the allowed values of k depend on the length of the lattice, $L = Na$. The lattice spacing a is not a quantization length for k. Its quantization length is L.

Interesting things happen when k assumes values close to either edge of the Brillouin zones $\pm \pi / a$. Referring to eqn (4.3), because k serves to define the eigenvalue λ_k, it is an appropriate index for the eigenfunctions and eigenvalues of T. Because a is real, k must also be real to ensure that $| \lambda_k | = 1$.

We shall forego the placement of a subscript on k to denote its different possible values. It is understood that k assumes different values, and that for a large lattice there is a nearly continuous range of k values. We shall see in due course that there can be many k values for each energy eigenvalue. However, these additional k values do not correspond to physically distinct states. In other words, there is mathematical redundancy. We shall also see that there are many energy eigenvalues for each k value. Later, when the eigenvalue spectrum is discussed, a second index (n) will be introduced to label states according to their atomic orbital parentages. For now, n is taken as understood.

Returning to the discrete translational symmetry, putting eqn (4.3) into eqn (4.2) gives

$$T\psi_k(x) = e^{ika}\psi_k(x). \tag{4.4}$$

It is now assumed that $\psi_k(x)$ is of the form

$$\psi_k(x) = u_k(x)e^{ikx}. \tag{4.5}$$

This form for $\psi_k(x)$ satisfies the disparate limits of free versus highly localized. To visualize these limiting cases, assume that the magnitude of some periodic potential $V(x)$ that localizes electrons at lattice sites is V_0. The waves e^{ikx} are appropriate for the limit $V_0 \rightarrow 0$. Here, $u_k(x)$ approaches a constant value, and k can assume an essentially continuous range of values, because it is quantized only via the overall length of the lattice, which is taken as huge. Thus, for $V_0 = 0$, the solutions of the Schrödinger equation are free particle waves, e^{ikx}, whose energies are $\hbar^2 k^2 / 2m$. The levels are degenerate for k values that have the same magnitude but differ in sign.

For $V_0 \neq 0$, the functions $u_k(x)$ conform to the lattice periodicity, namely, $|u_k(x)|^2$ does not vary from site to site. For small V_0 values, $u_k(x)$ can be said to modulate the plane waves. When V_0 is negative and has large magnitude, however, the functions $u_k(x)e^{ikx}$ are localized at the sites. With $u_k(x)$ localized to the extent that its amplitude is nearly zero between sites, there is little electron transport. An example of this is an atomic core level that hardly senses adjacent atoms.

Introducing the $\psi_k(x)$ given by eqn (4.5) into the right hand side of eqn (4.1) yields

$$T\psi_k(x) = u_k(x+a)e^{ikx}e^{ika}. \tag{4.6}$$

Likewise, putting this $\psi_k(x)$ into the right hand side of eqn (4.4) yields

$$T\psi_k(x) = e^{ika}u_k(x)e^{ikx} \tag{4.7}$$

$$= u_k(x)\,e^{ikx}e^{ika}\,. \tag{4.8}$$

A comparison of the right hand sides of eqns (4.6) and (4.8) indicates that

$$u_k(x+a) = u_k(x)\,. \tag{4.9}$$

The consistency that is achieved when eqn (4.5) is inserted into eqns (4.1) and (4.4), yielding eqn (4.9), verifies that $\psi_k(x)$ can be represented as:[2]

$$\boxed{\psi_k(x) = u_k(x)e^{ikx}\,.} \tag{4.10}$$

These are called Bloch waves, after Felix Bloch, who derived this equation and used it in crystal physics. It had been derived earlier in the field of mathematics, where it is referred to as Floquet's theorem. Figure 1 illustrates how it works by plotting the real part of $u_k(x)e^{ikx}$, with $u_k(x)$ taken to be real: $\mathrm{Re}\,\psi_k(x) = u_k(x)\cos kx$. You should figure out why it is acceptable to take $u_k(x)$ as real.

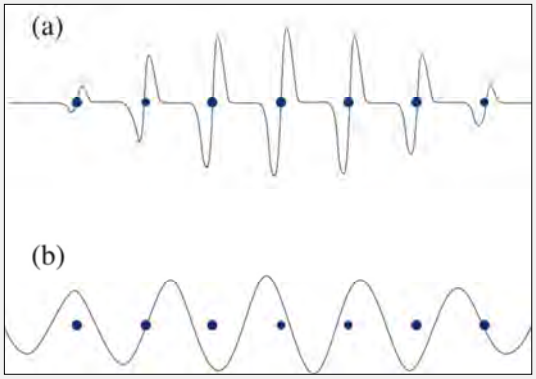


Figure 1. These plots show $\mathrm{Re}\,\psi_k(x) = u_k(x)\cos kx$ versus x. (a) $u_k(x)$ is proportional to $(x-x_i)\exp(-(x-x_i)^2)$, with lattice sites separated from one another by $|x_i - x_{i\pm1}| = 3$. There is essentially no electron transport. The long wavelength sinusoidal envelope is due to the fact that only $\mathrm{Re}\,\psi_k(x) = u_k(x)\cos kx$ is being plotted. The envelope of $\mathrm{Im}\,\psi_k(x)$ is out of phase with that of $\mathrm{Re}\,\psi_k(x)$. Of course $|\psi_k(x)| = |\psi_k(x+a)|$. In other words, peaks of $|\psi_k(x)|$ have constant site-to-site magnitude. (b) $u_k(x)$ is proportional to $\exp(-0.1(x-x_i)^2)$, so electrons are highly delocalized. The oscillations are due to $\cos kx$. In this case, the traveling wave can be said to modulate the $u_k(x)$. The value of k in (b) is larger than that in (a). Notice that the oscillations are not correlated with the positions of the nuclei. In this case, $|\psi_k(x)|$ is approximately constant throughout the lattice.

[2] We have not derived the Bloch form given in eqn (4.10) with mathematical rigor. It is unlikely that doing so would lead to additional enlightenment.

The traveling wave character of $\psi_k(x)$ is given by e^{ikx}. Using either the real or imaginary part of e^{ikx} (as long as $u_k(x)$ is real) enables the phase progression to be seen, whereas this information is lost with the modulus $|\psi_k(x)|$. In Fig. 1(a), the value of k is small enough that the variation of $\cos kx$ is slow relative to that of $u_k(x)$.

The probability density $|\psi_k(x)|^2$ is equal to $|u_k(x)|^2$, so it is the overlap of the various $u_k(x)$ with adjacent functions $u_k(x \pm a)$ that determines whether or not there is electron transport. If $u_k(x)$ is such that there is negligible overlap between adjacent sites, then there is essentially no net traveling-wave character, regardless of the e^{ikx} factor, and the electrons reside at the sites. Figure 1(a) is analogous to a core level: $\psi_k(x)$ goes almost to zero between sites, so there is little site-to-site electron transport. Another way to appreciate this is to note that the probability flux J is conserved in these systems. It can be evaluated anywhere along the x-axis. If it is evaluated where $\psi_k(x) = 0$, then $J = 0$ everywhere.

At the other extreme, when the electron is highly delocalized, the functions $u_k(x)$ have large amplitude between sites, even extending over a number of sites. In this case, we can think of e^{ikx} as modulating a relatively constant $u_k(x)$, as shown in Fig. 1(b).

Crystal Momentum

One should not think of k as momentum. It has been introduced as a parameter, and $\hbar k$ is called crystal momentum. Though the function e^{ikx} is associated with traveling wave character, it multiplies $u_k(x)$. Thus, $\hbar k$ is not a momentum in the sense of a free particle, where $p = \hbar k$ applies. The linear momentum of a free particle is conserved because of the homogeneity of free space. The displacement operator applies to an infinitesimal displacement anywhere in free space, and consequently the translational invariance of p arises from a continuous symmetry. On the other hand, the T we have introduced for site-to-site displacement is a discrete symmetry. Therefore, its eigenfunctions are not eigenfunctions of p. To see how this works mathematically, let p operate on $\psi_k(x)$:

$$p\,\psi_k(x) = -i\hbar\nabla\left(u_k(x)e^{ikx}\right) \tag{4.11}$$

$$= \hbar k\psi_k(x) - i\hbar\, e^{ikx}\,\nabla u_k(x)\,. \tag{4.12}$$

Thus, $p\,\psi_k(x)$ is not equal to $\psi_k(x)$ times a constant, and therefore $\psi_k(x)$ is not an eigenfunction of p. When localization at the sites is weak, $\nabla u_k(x)$ is small and $p\,\psi_k(x)$ approaches $\hbar k\,\psi_k(x)$. Likewise, in this case the group velocity: $v_g = \partial E / \partial \hbar k$, approaches $\hbar k / m$. This is the nearly free electron regime. At the other extreme, when $u_k(x)$ is highly localized, there is little overlap between $u_k(x)$ and $u_k(x+a)$, as seen in Fig. 1(a). In this case, there is little site-to-site flux because $\psi_k(x)$ nearly vanishes between sites. This is the tight-binding regime. Though $\hbar k$ is not a momentum eigenfunction, it has units of momentum, and it is the momentum in the free electron limit. Therefore, we need to figure out just what kind of momentum it might be.

The relationship between $\hbar k$ and momentum can be determined by considering the affect of an applied electric field. Bloch waves are time independent solutions to the Schrö-

dinger equation. They reflect the fact that in a perfect crystal with stationary nuclei there is no resistance to the flow of electric current. In other words, the ansatz that the nuclei are fixed in space means that no phonons can be created. Thus, to see what kind of momentum $\hbar k$ might be, a quasiclassical situation is constructed in which a wave packet comprising Bloch waves moves through the crystal under the influence of the electric field $\vec{E}$ without being scattered by phonons.

The energy of the electron increases as it gains energy from the field, and, because there is no loss, the sum of its translational and potential energies is constant:

$$E_0 - e\phi = \text{constant}\,, \tag{4.13}$$

where E_0 is the center translational energy of the wave packet, and ϕ is the electric potential. Differentiation yields

$$\dot{E}_0 - e\dot{\phi} = 0\,. \tag{4.14}$$

Minor manipulation gives

$$\frac{\partial E_0}{\partial k}\dot{k} - e\left(\nabla\phi\right)\dot{x} = 0. \tag{4.15}$$

In going from eqn (4.14) to eqn (4.15), use is made of the fact that, for a constant electric field, E_0 can be expressed as an explicit function of k, and ϕ can be expressed as an explicit function of x. Next, use is made of the facts that: (i) $\vec{E} = -\nabla\phi$ (the $-\partial_t \vec{A}$ term is neglected in this quasi-static regime); (ii) the group velocity v_g is equal to $\partial E_0 / \partial \hbar k$ (Appendix 2); and (iii) $\dot{x}$ is equal to v_g in the present case. Thus, eqn (4.15) becomes

$$\left(\hbar\dot{k} + e\vec{E}\right)v_g = 0. \tag{4.16}$$

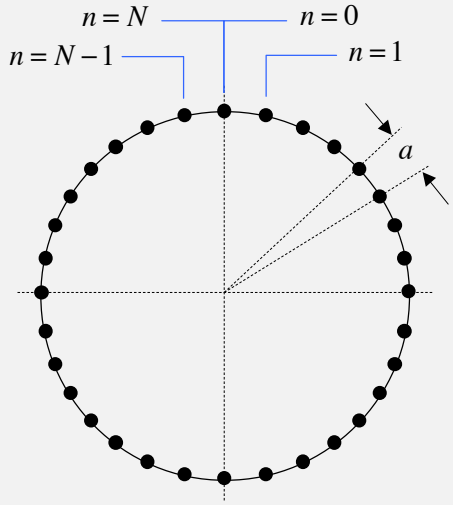
This shows that $\hbar k$ responds as a momentum to the applied electric field. The distinction between $\hbar k$ and p is clarified. Momentum p responds to all forces, including those from the binding sites. The crystal momentum behaves as a translational momentum, but to an applied field that causes current to flow. Its time rate of change $\hbar\dot{k}$ is equal to the force $\vec{F} = q\vec{E} = -e\vec{E}$. For core levels, there is essentially no electron transport, so $v_g \to 0$, in which case ($\hbar\dot{k} + e\vec{E}$) and v_g in eqn (4.16) vanishe separately.

Allowed k Values

The allowed values of k are obtained by considering a lattice of length $L = Na$, where N is the number of sites. As in Chapters 1 – 3, the ends are treated mathematically as being connected to one another, in which case the lattice is modeled as a giant ring, as indicated in Fig. 2. You should think of this figure as a mnemonic for imposing periodic boundary conditions. The requirement that the wave function is continuous where the ends of the 1D lattice join together demands

$$\psi_k(x) = \psi_k(x + Na). \tag{4.17}$$

Figure 2. A one-dimensional lattice is wrapped into a giant ring: $Na = L$. In this case, the boundary condition on the wave function is $\psi_k(x) = \psi_k(x + Na)$. The wave function's phase varies with location on the circle. It can be expressed in terms of the phase progression between adjacent sites, e^{ika}.



Equation (4.17) is now used with eqn (4.2). Namely, applying the T operator N times in succession carries x to $x + Na$, and eqn (4.17) requires:

$$T^N \psi_k(x) = \lambda_k{}^N \psi_k(x) \tag{4.18}$$

$$= \psi_k(x + Na) \tag{4.19}$$

$$= \psi_k(x). \tag{4.20}$$

Thus, $\lambda_k{}^N = 1$, which has N roots, each with unit magnitude and a phase. The eigenvalues λ_k are expressed as

$$\lambda_k = \exp\left(i\frac{2\pi}{N}m\right) \quad \text{where } m = 0, 1, 2 \ldots N-1. \tag{4.21}$$

The integer m assumes N sequential values. For example, this range can be 0, 1, 2 … $N-1$ (as shown above). Alternatively, for odd N it could be: $-(N-1)/2 \ldots 0 \ldots (N-1)/2$, whereas for even N it could be: $-N/2 \ldots 0 \ldots N/2$, with the understanding that $-N/2$ and $N/2$ contribute one state, not two. Combining eqns (4.1) and (4.2), and using $\psi_k(x) = u_k(x)e^{ikx}$ and the expression for λ_k given by eqn (4.21) yields

$$\psi_k(x+a) = e^{i2\pi m/N}u_k(x)e^{ikx}, \tag{4.22}$$

and using $u_k(x) = u_k(x+a)$ gives

$$\psi_k(x+a) = u_k(x+a)e^{i(kx+2\pi m/N)}. \tag{4.23}$$

The right hand side of eqn (4.23) is equal to $u_k(x+a)e^{ik(x+a)}$ only if

$$k = \frac{2\pi}{Na}m = \frac{2\pi}{L}m, \tag{4.24}$$

where L is the circumference of the large ring. Make sure you understand how this stuff works before moving on.

Equation (4.24) is the quantization condition. The system is doubly degenerate: $|\pm m|$, with + and – corresponding to opposite senses of circulation. The smallest k values have wavelengths that are much larger than the lattice spacing. As k increases, the wavelength eventually becomes comparable to the lattice spacing. Though the large-L limit yields the infinite lattice, the Bloch theorem works for any value of $L = Na$.

The k that has been introduced is a parameter in the phase evolution factors e^{ikx} in the Bloch functions, and it labels the Bloch functions. It enters via the discrete symmetry of displacement by the lattice spacing a. As such, it is not related to the momentum p except when a is unimportant, that is, $u_k(x) \rightarrow$ a constant. Recall that in Chapter 1 it was seen that in the limit of a weak potential a periodic potential couples states such that gaps appear when $m' + m \sim 0$.

4.2. Brillouin Zones

Let us now examine the fact that the energy of a Bloch function is unaffected if k is increased or decreased by an integer multiple of $2\pi/a$. We shall, of course, invoke the system's discrete translational symmetry, which ensures that simultaneous eigenfunctions of T and H can be obtained. This symmetry lies at the heart of the matter, by definition, as it is the only spatial symmetry in the model. To begin, note that e^{ikx} undergoes no change from one site to the next when k is replaced by $k+G$, where the 1D reciprocal lattice vector, G, is given by $2\pi m_G/a$, where m_G is an integer.[3] Namely, phase progression from one site to the next is e^{ika}, and replacing k with $k+G$ yields

$$e^{i(k+G)a} = e^{ika}e^{iGa} = e^{ika}e^{i2\pi m_G} = e^{ika}. \tag{4.25}$$

Thus, $iGa = i2\pi m_G$ can be added to the argument of the exponential with impunity, and $T\psi_k(x)$ can be written:

$$T\psi_k(x) = e^{ika}\psi_k(x) = e^{i(k+G)a}\psi_k(x). \tag{4.26}$$

This shows that adding G to the k in e^{ika} does not change the state. Of course, if T operating on a state gives the eigenvalue $e^{i(k+G)a}$, the state must be $\psi_{k+G}(x)$. Therefore $\psi_k(x)$ must be equal to $\psi_{k+G}(x)$. Alternatively, one can start with $\psi_k(x) = e^{ikx}u_k(x)$ and multiply e^{ikx} by e^{iGx} and $u_k(x)$ by e^{-iGx}. This leaves $\psi_k(x)$ unaffected:

$$\psi_k(x) = e^{i(k+G)x}\left(e^{-iGx}u_k(x)\right). \tag{4.27}$$

The term in large parentheses is periodic in a, so the right hand side is $\psi_{k+G}(x)$. This means that E_k is equal to E_{k+G}, as k and $k+G$ each denote the same state.[4]

A mathematical redundancy has been revealed. For a given value of E, there can exist many values of $k+G$. This is best illustrated with an example. Figure 3 uses results from Fig. 7 of Chapter 1 for the case of 24 sites, with smooth curves replacing the 24 discrete

---

[3] The subscript G on m_G is used to distinguish this integer from the one that appears in the quantization of k according to the length of the ring: $k = 2\pi m/L$.

[4] Yet another perspective considers the Bloch function $\psi_{k+G}(x) = e^{ikx}e^{iGx}u_{k+G}(x)$. Because $Ga = 2\pi m_G$ and $u_{k+G}(x) = u_{k+G}(x+a)$, this can be written

$$\psi_{k+G}(x) = e^{ikx}\left(e^{iG(x+a)}u_{k+G}(x+a)\right).$$

The term in large parentheses, being periodic in a, behaves the same way as $u_k(x)$. Thus, $\psi_{k+G}(x)$ and $\psi_k(x)$ have the same properties. When a value of E is found that corresponds to a specific value of k, it follows that $2\pi m_G/a$ ($m_G = \pm1, \pm2...$) can be added to this k without changing E.

points per band. Referring to Fig. 3(a), notice that the first Brillouin zone ($-\pi/a \le k \le \pi/a$) is repeated *ad infinitum* with increasing $|k|$. This is a direct consequence of the Bloch theorem, as noted with eqns (4.25) and (4.26). Referring to Fig. 3(b), there are 24 levels in the interval: $-\pi/a \le k \le \pi/a$. These were obtained in Fig. 7 of Chapter 1 for a weak binding potential by using matrix diagonalization and a particle-on-a-ring basis.

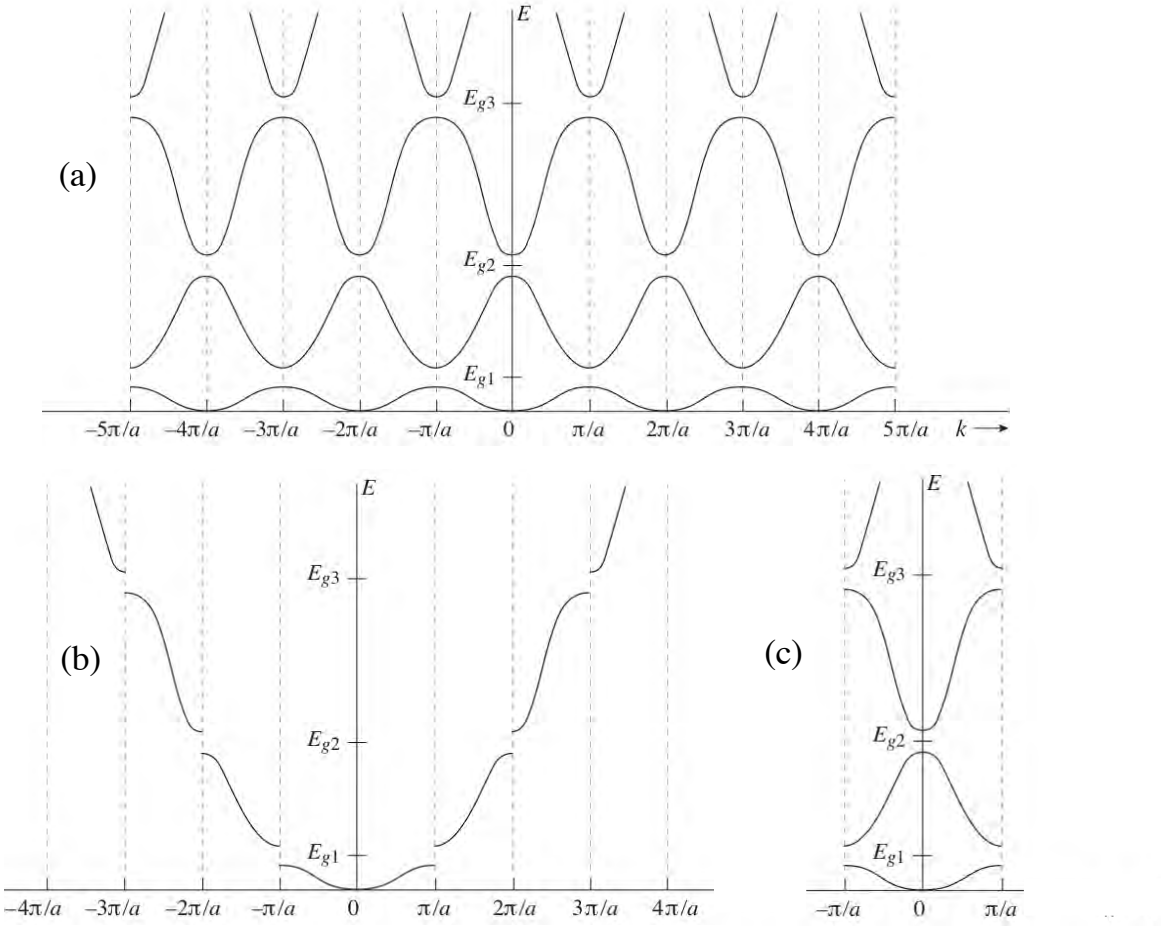


Figure 3. (a) For a given E within a band, there exist two k values within a $2\pi/a$ interval, as well as k values displaced from these by integer multiples of $2\pi/a$. All such possibilities are shown in the repeated zone representation. Only one $2\pi/a$ interval is unique and has physical meaning. (b) For the nearly free electron, E varies quadratically with k except near the gaps. The extended zone representation displays this. In (c), the restricted zone representation places all of the information in the first Brillouin zone.

The important point is that, for a given band, it is only possible to have 24 molecular orbitals because there are only 24 binding sites. For example, consider the lowest band in Fig. 3(b). When k fills the interval $-\pi/a \le k \le \pi/a$, that is it. Values of k outside this range (with $E < E_{g1}$) simply repeat states that are already present in the original 24. This redundancy is explicit in Fig. 3(a). For each band of allowed energies the shape within a single Brillouin zone is repeated in the horizontal directions.

At the same time, there must also be cases in which the wave vector $k + G$ belongs to a state that is distinct from ψ_k. For example, in the free electron limit, energies are given by $\hbar^2 k^2 / 2m$. Therefore, a state must exist whose energy is $\hbar^2 (k+G)^2 / 2m$. This state can be written: $e^{i(k+G)x} = e^{ikx} u_k(x)$, where $u_k(x) = e^{iGx}$. We see that it is then straightforward to include higher momentum in which $G = 2\pi m_G / a$. As with the lowest band, a length of $2\pi / a$ in *k*-space contains 24 states. The rest are repeats. In Fig. 3(b) the length of $2\pi / a$ for the next highest band is the combined intervals π / a to $2\pi / a$ and $-\pi / a$ to $-2\pi / a$.

To summarize, the information that constitutes the 1D band structure can be viewed in different ways: fill the paper with repeating segments; assign the segments to physically meaningful portions of *E* versus *k*; or line up the bands vertically in the first Brillouin zone. When the bands are filled (no electron transport), the restricted zone representation shown in Fig. 3(c) is intuitive. When spin ½ is included, the number of states in a band is equal to twice the number of sites. Thus, in simple terms, when each site donates one electron to a band, the band is half full. When each site donates two electrons to a band, the band is full. When the energy gaps between the bands are large, electrons do not move between the bands to any significant extent. Figure 3(a) is of course correct, but it is too redundant to be practical. In (a) – (c), there is symmetry about the midpoint so only half of the Brillouin zones need to be shown. Keep in mind that Fig. 3 is a toy model in the extreme. It is not meant to mimic a real system.

4.3. Energy Eigenvalues

The above results can be viewed from the perspective of the energy eigenvalues by putting $\psi_k(x)$ into the Schrödinger equation

$$\left(\frac{p^2}{2m} + V(x) \right) \psi_k(x) = E(k) \psi_k(x) . \tag{4.28}$$

In eqn (4.28), $E(k)$ is used instead of E_k because we are interested here in a lattice consisting of many sites, in which case *k* varies almost continuously, quantized only by the length $L = Na$. Note that $E(k)$ is not a density. It has units of energy, not energy per unit wave number, as would be the case for a density. Introducing $\psi_k(x) = e^{ikx} u_k(x)$, and following some algebra, eqn (4.28) becomes

$$\left(-\frac{\hbar^2}{2m} (ik + \nabla)^2 + V(x) \right) u_k(x) = E(k) u_k(x) . \tag{4.29}$$

Notice that the exponential factor e^{ikx} is no longer present. From the presence of *k* in the term $(ik + \nabla)$, we see why *k* is an appropriate label for Bloch functions. Equation (4.29) is a Hermitian eigenvalue equation, so for each value of *k*, there exists a spectrum of real eigenvalues. These can be labeled $E_n(k)$, where *n* indexes the eigenvalues of eqn

(4.29) for a given k. Thus, the Bloch function given by eqn (4.10) is now rewritten to include the index n:

$$\psi_{k,n}(x) = u_{k,n}(x)e^{ikx}. \tag{4.30}$$

The Bloch functions $\psi_{k,n}(x)$ are such that, for each value of k, a large number of eigenfunctions are labeled using the index n. Each different n has parentage in an atomic orbital. You might find it useful to think of n as a principal quantum number. Mathematically, the maximum n value can approach infinity, but in practice it is limited to orbitals without extreme excitations.

As mentioned above, all energy eigenvalues are contained in a range of ka values that spans 2π. The range $-\pi/a \le k \le \pi/a$ gets them all. Letting $k \to k + 2\pi m_G/a$, where m_G is an integer, does not change E. The region bounded by $-\pi/a \le k \le \pi/a$ is the first Brillouin zone. The second Brillouin zone is contained in the regions $-2\pi/a \le k \le -\pi/a$ and $\pi/a \le k \le 2\pi/a$. And so on for higher Brillouin zones.

Now return to Fig. 3(c), choose any k in the first Brillouin zone, and draw a vertical line through this k. The intersection points with the excited curves are the eigenvalues $E_n(k)$. This is the meaning of the Hermitian eigenvalue problem represented with eqn (4.29). You can also locate these solutions on Fig. 3(b) by shifting k by integer multiples of $2\pi/a$ and drawing vertical lines.

Fourier Analysis

A complementary way to view the eigenvalue problem enlists Fourier analysis. This focuses on the relationship between three things: wave vector k, the shape of the potential $V(x)$, and the eigenvalues. Let us expand the wave function in a plane wave basis that is periodic in the length of the lattice (Fig. 4). This is how the particle-on-a-ring model of Chapter 1 was approached. There, the quantization length was that of the large ring, with periodic boundary conditions, and we chose the basis $e^{im\phi}$. It was seen that when $m - m'$ was commensurate with the potential's periodicity the coupling matrix elements were nonzero.

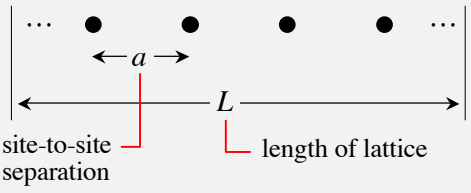


Fig. 4. The wave function quantization length is that of the lattice. The fundamental wave vector is $2\pi/L$ for periodic boundary conditions. The potential is expanded within one repeating interval. Its fundamental wave vector is $2\pi/a$.

The potential $V(x)$ shall be taken as symmetric about $x = 0$. It is expanded in terms of reciprocal lattice vectors G. This is another way of saying that its Fourier expansion consists of harmonics defined by the properties of an individual site. You should make a few sketches to see how this works. Fourier expansions for the wave function and potential can be written

$$\psi(x) = \sum_k C_k \, e^{ikx} \tag{4.31}$$

$$V(x) = \sum_{G>0} 2V_G \cos Gx \tag{4.32}$$

$$= \sum_{all\,G} V_G \, e^{iGx} \, . \tag{4.33}$$

Keep in mind that the $\psi(x)$ in eqn (4.31) is not written in the Bloch form given by eqn (4.30). Rather, it is expanded in a plane wave basis. The G's in eqn (4.33) have positive and negative values, as they must in order to ensure that $V(x)$ is real. The $G = 0$ term in the sum is excluded because it is an unimportant offset. In other words, summation over "all G" in eqn (4.33) means all positive and negative integer multiples of $2\pi / a$, excluding zero. In eqn (4.31), the smallest possible k value is $2\pi / L$ (periodic boundary condition), and the sum is over integer multiples of this. In eqn (4.33), the smallest possible G value is $2\pi / a$, and the sum is over integer multiples of this.

Putting eqns (4.31) and (4.33) into the Schrödinger equation will yield, in a few steps, conditions for the coefficients C_k. To start, the Schrödinger equation becomes

$$\frac{\hbar^2}{2m} \sum_k k^2 C_k e^{ikx} + \sum_k \sum_G V_G C_k e^{i(k+G)x} = E \sum_k C_k e^{ikx} \, . \tag{4.34}$$

The sum over k in the second term can be rewritten using the substitution: $k' = k + G$. The sum over k is then replaced by a sum over k'. Thus, C_k becomes $C_{k'-G}$, and the argument in the exponent is $ik'x$. Because the sum is over all k' (addition of G has the effect of shifting the origin) the prime is dropped, and eqn (4.34) reads

$$\sum_k \left(\left(\frac{\hbar^2 k^2}{2m} - E \right) C_k + \sum_G V_G C_{k-G} \right) e^{ikx} = 0 \, . \tag{4.35}$$

Equation eqn (4.35) holds for each value of k. Thus, we have

$$\boxed{\left(\frac{\hbar^2 k^2}{2m} - E \right) C_k + \sum_G V_G C_{k-G} = 0 \, .} \tag{4.36}$$

Equation (4.36) shows that, for a given k value, the only coefficients in the expansion of the wave function are those for which k changes by an integer multiple of the fundamental reciprocal lattice vector $2\pi / a$, as shown in Fig. 5. This result was obtained from a different perspective in eqn (4.27), where $\psi_{k+G}(x)$ was shown to have the same energy spectrum as $\psi_k(x)$. Thus, each value of E corresponds to all values of $k + G$,

where G is an integer multiple of $2\pi/a$. You should think about this subtle point for a while before moving on.

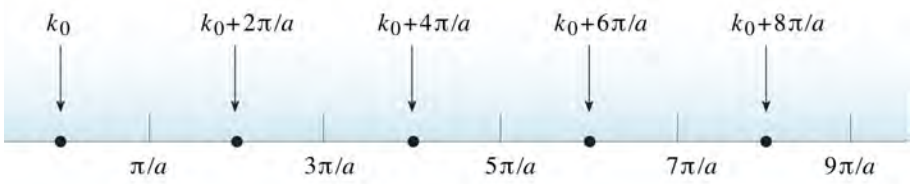


Figure 5. The system is unchanged when an integer multiple of $2\pi/a$ is added to any value of k. For example, k_0 can be changed to $k_0 + 2\pi/a$, $k_0 + 4\pi/a$, and so on.

Equation (4.36) of course is consistent with the fact that as the potential goes to zero E approaches the free particle limit $\hbar^2k^2/2m$. Alternatively, if the potential is sinusoidal with period a, for example, $V_0\cos(2\pi x/a)$, the sum contributes just two terms, corresponding to $\pm|G| = \pm 2\pi/a$. All other V_G values are equal to zero. The percentages of the higher harmonics in the expansion of the potential given by eqn (4.33) dictates the range of C_{k-G} values needed to express the eigenfunctions.

The result we have just obtained might strike you as formal. However, we encountered it earlier without the mathematical development that has been used here. Go back to Figs. 1.7 – 1.12 and recall the discussions accompanying these diagrams: harmonics of $V(x)$, avoided crossings, resonance conditions, and so on. It is more or less the same material.

Zone Boundary

Now consider a potential $V(x)$ that is sufficiently weak that plane waves are good descriptors except near the band edges. In particular, consider the edge of the first Brillouin zone where $k = \pm\pi/a$. In this case a half wavelength is equal to the site separation a, and $e^{i\pi x/a}$ and $e^{-i\pi x/a}$ combine to form sines and cosines, one peaked where the potential is maximum, the other where it is minimum (Fig. 6). This explains the band gap, which is the energy between these very different states. The numerical value of the band

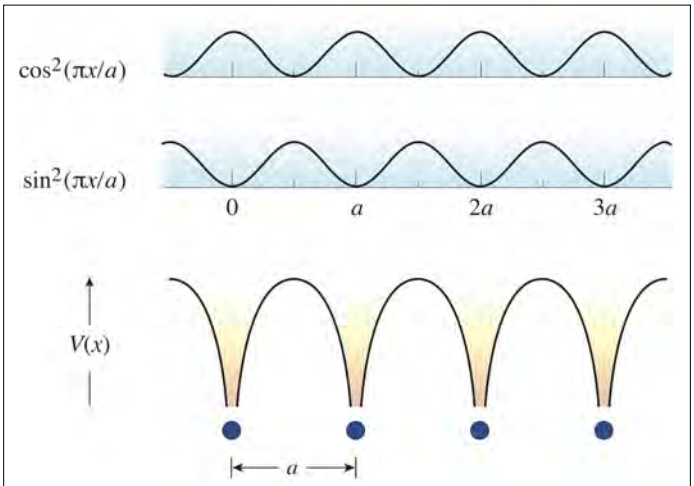


Figure 6. An electron experiences a weak periodic potential due to ionized nuclei separated by a. Standing waves are formed at $k = \pm\pi/a$.

gap is obtained by averaging $V(x)$ over the difference between the probability densities that correspond to the sine and cosine solutions:

$$E_g = \frac{2}{a}\int_0^a dx\left(\cos^2\left(\frac{\pi x}{a}\right)-\sin^2\left(\frac{\pi x}{a}\right)\right)V(x). \tag{4.37}$$

The factor $2/a$ normalizes the wave functions over the interval a. The trigonometric identity: $\cos^2 \mathrm{A} - \sin^2 \mathrm{A} = \cos 2\mathrm{A}$, yields

$$E_g = \frac{2}{a}\int_0^a dx\cos\left(\frac{2\pi x}{a}\right)V(x). \tag{4.38}$$

The cosine term contains the periodicity of the difference between the upper and lower level probability densities. The potential $V(x)$ is expanded in a Fourier series. Only the term in the Fourier series that varies as $\cos(2\pi x/a)$ (the lowest order term in the expansion) survives integration. Equation (4.38) yields the band gap energy at the edge of the first Brillouin zone.

Two limiting cases are now examined: (i) binding is weak enough that, to a first approximation, the electron can be taken as nearly free; and (ii) binding is strong enough that, to a first approximation, site-to-site electron transport is quite modest.

Exercise: Explain why eqns (4.36) and (4.37) give the band gap. Should we not be taking the difference between eigenstates or between expectation values of the Hamiltonian?

4.4. Nearly Free Electron

Suppose an electron in a 1D lattice comprising many binding sites somehow managed to be oblivious to the sites. In this case, its motion would be restricted only in the sense that the electron wave function must respect the boundary condition imposed by the overall length of the lattice. The electron would have wave functions e^{ikx} and energies $\hbar^2k^2/2m$. Momentum would be quantized according to the length of the lattice, which is large, yielding the parabolic $E(k)$ indicated in Fig. 7. On the other hand, if the electron experiences a weak periodic potential, its eigenfunctions must be Bloch waves $\psi_{k,n}(x) = u_{k,n}(x)e^{ikx}$, where $u_{k,n}(x)$ is nearly constant. This case of the so-called nearly free electron is treated here. The electron is referred to as nearly free because there exists a lattice potential that, however slight, influences electrons having k values in the vicinities of $\pm\pi/a$, $\pm 2\pi/a$, and so on.

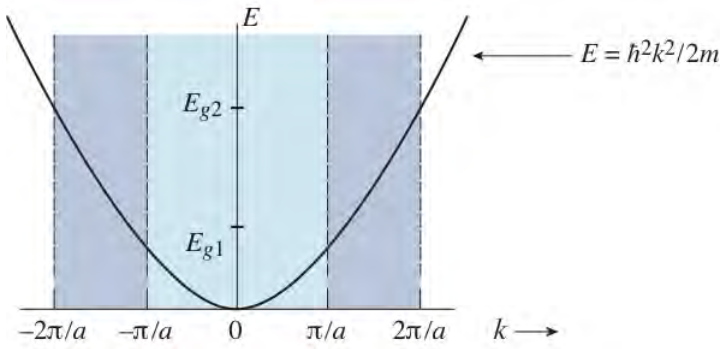


Figure 7. The $E(k)$ parabola is the dispersion relation for a free electron. In this limit, the sites and Brillouin zones are irrelevant.

For k values well away from the edges of the Brillouin zones, the e^{ikx} waves are good descriptors. At the Brillouin zone edges, however, strong effects arise from even weak interaction with the lattice. In the vicinities of the Brillouin zone edges, states whose k values differ by integer multiples of $2\pi/a$ (with $k \approx |k'|$) are coupled strongly to one another. The interpretation $p = \hbar k$ cannot be used, and energy is no longer given by $\hbar^2k^2/2m$.

For example, we have seen that at the edge of the first Brillouin zone, the degenerate pair $e^{i\pi x/a}$ and $e^{-i\pi x/a}$ yield standing waves in which the upper and lower states correspond to probability densities peaked in regions of high and low potential, respectively. Applied to the other Brillouin zone boundaries, small gaps appear in the $E(k)$ curves when $k = \pm 2\pi/a, \pm 3\pi/a$, and so on, as indicated schematically in Fig. 8.

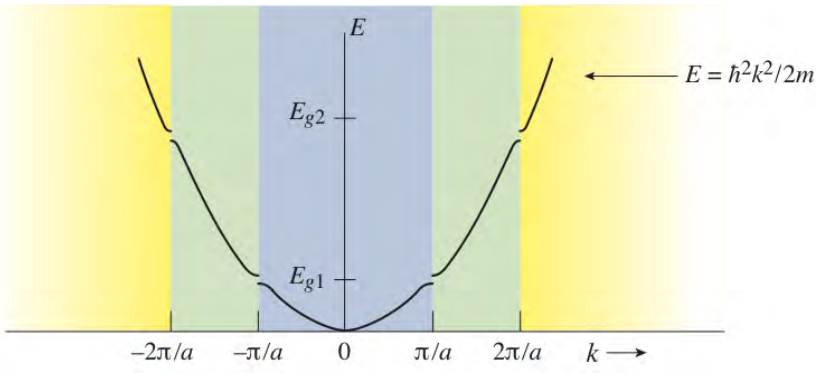


Figure 8. The nearly free electron model: small gaps appear at k values of $\pm \pi/a, \pm 2\pi/a$, and so on: the edges of the Brillouin zones.

The interactions at the Brillouin zone edges are curve crossings, as we saw way back in Chapter 1. Referring to Fig. 8, translating by $2\pi/a$ the region near $k = -\pi/a$ to the region near $k = \pi/a$ results in the crossing indicated in Fig. 9. In its vicinity, it is easy to evaluate the coupling matrix element in the traveling wave basis, and the eigenvalues are readily obtained using a 2×2 matrix. There is no need to consider influences brought about by interactions with higher levels, because of the fact that we are at a region of near-degeneracy and coupling is weak.

The coupling matrix element that is responsible for the energy gap at the edge of the first Brillouin zone is now evaluated between levels k in the first Brillouin zone and $k - 2\pi/a$ in the second Brillouin zone. In the plane wave basis, this matrix element is

$$\left\langle k \left| V(x) \right| k - \frac{2\pi}{a} \right\rangle = \frac{1}{L} \int_0^L dx\, e^{-i2\pi x/a}\, V(x) . \tag{4.39}$$

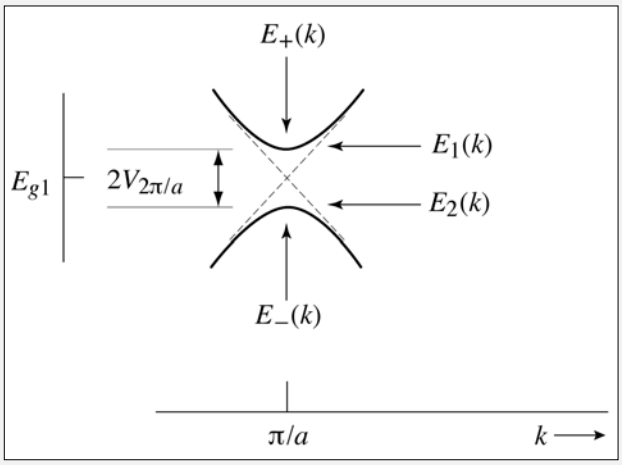


Figure 9. In the vicinity of the curve crossing, only $E_1(k)$ and $E_2(k)$ (dashed lines) need to be considered. Plane waves interact via the periodic potential, creating a gap: $2|V_{2\pi/a}|$ at $k = \pi/a$. The $E_+(k)$ and $E_-(k)$ curves have parabolic shapes in the vicinity of $k = \pi/a$.

The integral retains only the leading term in the Fourier expansion of $V(x)$, which is expanded in harmonics of the fundamental reciprocal lattice vector $2\pi/a$. In short, integration retains only the $2\pi/a$ term in the expansion of $V(x)$. Diagonal matrix elements are zero. Earlier, the Fourier expansion of the periodic potential was introduced using sine, cosine, and complex exponential basis functions whose arguments were iGx, with G given by integer multiples of $2\pi/a$.[5] With $V(x)$ expressed in the exponential basis, eqn (4.39) becomes:

$$\left\langle k \left| V(x) \right| k - \frac{2\pi}{a} \right\rangle = \frac{1}{L} \int_0^L dx \left(e^{-i2\pi x/a} \sum_G V_G\, e^{iGx} \right) \tag{4.40}$$

$$= \sum_G V_G \frac{1}{L} \int_0^L dx \exp\left(-i \left(\frac{2\pi}{a} - G \right) x \right) . \tag{4.41}$$

As mentioned above, the only non-vanishing integration in the sum over all G is the one for which $G = 2\pi/a$. Thus, eqn (4.41) gives a simple and familiar result:

[5] Note that there is a difference of a factor of two in the magnitudes of the expansion coefficients for real versus complex exponential bases.

$$\left\langle k \left| V(x) \right| k - \frac{2\pi}{a} \right\rangle = V_{2\pi/a} . \tag{4.42}$$

The off-diagonal matrix element $V_{2\pi/a}$ is now used with the 2×2 Hamiltonian matrix that describes coupling in the vicinity of the band gap. Referring to Fig. 9, the eigenvalues of the 2×2 matrix are

$$E_{\pm} = E_{g1} \pm \sqrt{\tfrac{1}{4}\left(E_1(k) - E_2(k)\right)^2 + |V_{2\pi/a}|^2} , \tag{4.43}$$

where E_{g1} is the crossing point indicated in Fig. 9, and $E_1(k)$ and $E_2(k)$ are the undisturbed curves, that is, the dashed lines in Fig. 9. Except for the continuous nature of $E_1(k)$ and $E_2(k)$, eqn (4.43) is the same as one encountered in the Chapter 1 section entitled: *Pairwise Interaction*. The curves $E_1(k)$ and $E_2(k)$ are now expanded in a Taylor series about the crossing point at $k = \pi / a$. Referring to Fig. 9, retaining the lowest order (linear) term in the expansion gives

$$E_{1,2}(k) = E_{g1} \pm \left(\frac{\partial E}{\partial k} \right)_{\pi/a} \left(k - \frac{\pi}{a} \right) \tag{4.44}$$

$$= E_{g1} \pm \frac{\hbar^2}{m} \frac{\pi}{a} \left(k - \frac{\pi}{a} \right), \tag{4.45}$$

where $E = \hbar^2 k^2 / 2m$ has been used. Thus, the first term inside the radical in eqn (4.43) is given by

$$\tfrac{1}{4}\left(E_1(k) - E_2(k)\right)^2 = \left(\frac{\hbar^2}{2m} \frac{2\pi}{a} \left(k - \frac{\pi}{a} \right) \right)^2 \tag{4.46}$$

$$= 4 \left(\frac{\hbar^2 k'^2}{2m} \right) \left(\frac{\hbar^2 (\pi / a)^2}{2m} \right) \tag{4.47}$$

$$= \left(\frac{\hbar^2 k'^2}{2m} \right) 4 E_{g1} . \tag{4.48}$$

where $E_{g1} = \hbar^2 (\pi / a)^2 / 2m$ has been used, and the definition $k' = k - \pi / a$ has been introduced. Introducing this into eqn (4.43) yields

$$E_{\pm}(k') = E_{g1} \pm \sqrt{\frac{\hbar^2 k'^2}{2m} 4 E_{g1} + |V_{2\pi/a}|^2} . \tag{4.49}$$

Near $k' = 0$, the $|V_{2\pi/a}|^2$ term in the radical dominates. Expanding the radical, which is of the form $|V_{2\pi/a}|(1+\varepsilon)^{1/2}$, yields

$$E_{\pm}(k') \approx E_{g1} \pm |V_{2\pi/a}| \pm \alpha \frac{\hbar^2 k'^2}{2m}, \tag{4.50}$$

where $\alpha = 2E_{g1}/|V_{2\pi/a}|$ and $k' = k - \pi/a$.

Equation (4.50) shows that the $E_{\pm}(k')$ curves are parabolic near $k' = 0$. This is suggestive of states whose energies are "kinetic." However, given that the $k' = 0$ solutions are standing waves, the role of kinetic energy is curious. If one insists on assigning the k'^2 term to kinetic energy, it must be of the mathematical form $\hbar^2 k'^2/2m$. Moreover, to be consistent with eqn (4.50), the m in this expression must be replaced by a parameter. This parameter is referred to as the effective mass, and it is labeled m^*. The kinetic energy term in eqn (4.50) is then $\hbar^2 k'^2/2m^*$. Note that the $\pm$ that accompanies the term $\alpha\hbar^2 k'^2/2m$ has been absorbed into the effective mass, which therefore can be positive or negative.

The effective masses at the zone edges are defined from the curvatures of the $E_{\pm}(k')$ curves near their stationary points. Referring to eqn (4.50), at the edge of the first Brillouin zone, this gives

$$\left(\frac{\partial^2 E_{\pm}(k')}{\partial k'^2}\right)_0 = \pm\frac{\hbar^2\alpha}{m}. \tag{4.51}$$

This expression has been obtained using the nearly free electron model. The assignment: $\pm\hbar^2\alpha/m = \hbar^2/m^*$ is the definition of effective mass m^*. For the $E_{\pm}(k')$ curves described by eqn (4.50) and Fig. 9 we have

$$E_{\pm}(k') = \left(E_{g1} \pm |V_{2\pi/a}|\right) + \frac{\hbar^2}{2m^*}\left(k - \frac{\pi}{a}\right)^2. \tag{4.52}$$

($\pm m/\alpha$; $2E_{g1}/|V_{2\pi/a}|$)

In the nearly free electron limit, the energy of the band gap $2|V_{2\pi/a}|$ is much smaller than the energy E_{g1} that locates the center of the gap. Thus, $\alpha = 2E_{g1}/|V_{2\pi/a}|$ is much larger than unity, and the electron's effective mass is much smaller than its real mass. This is due to the lattice. It appears to someone oblivious to the lattice potential that the electrons are lighter than is in fact the case. The effect brought about through the potential has been parameterized into the effective mass.

4.5. Tight Binding

The limit opposite that of the nearly free electron is referred to as tight binding. The electron, though strongly bound to a site, is not so strongly bound that it completely loses communication with adjacent sites. For example, it is assumed that the electron resides at a given site for long periods of time, tunneling infrequently to adjacent sites.

When electrons in a given atomic orbital are bound strongly to the sites, the function $u_{k,n}(x)$ and its counterparts at the immediately adjacent sites, $u_{k,n}(x \pm a)$, have only modest spatial overlap. The atomic orbitals are good first approximations, and the energy bands are narrow. This applies to the low energy filled bands that comprise atomic core orbitals and, to a lesser extent, to the highest occupied bands of insulators. For a given atomic orbital that we shall call ϕ_ν, the Bloch function can be written

$$\psi_k(x) = \left(N^{-1/2} \sum_{i=1}^{N} e^{-ik(x-x_i)} \phi_\nu(x - x_i) \right) e^{ikx} . \tag{4.53}$$

In eqn (4.53), $x_i = ia$, where i is an integer that is not to be confused with the imaginary unit i. The subscript ν denotes the atomic orbital under consideration. The letter ν is used because it is the Greek counterpart to n. It is assumed that the atomic orbitals $\phi_\nu(x - x_i)$ are normalized. The index n for the corresponding band is suppressed in favor of a compact notation. It can be retrieved later. Thus, we write $\psi_k(x)$ instead of $\psi_{k,n}(x)$. The normalization factor $N^{-1/2}$ is an approximation based on the fact that electrons are tightly bound to the sites. It neglects the relatively small contribution due to overlap between wave functions centered at adjacent sites.

The large parenthetic term in eqn (4.53) is periodic in a, and therefore when it is multiplied by e^{ikx} we have a Bloch function. The electron atomic orbitals $\phi_\nu(x - x_i)$ are centered at the ion cores located at x_i. The phase evolution factor e^{ikx} on the right of the large parentheses does not imply electron transport, though it accommodates it when binding is not too tight. The tight binding limit, as its name implies, does not accommodate electron transport to any significant extent. This is made clear by writing eqn (4.53) as

$$\psi_k(x) = N^{-1/2} \sum_{i=1}^{N} e^{ikx_i} \, \phi_\nu(x - x_i) . \tag{4.54}$$

Equation (4.54) shows that $\psi_k(x)$ can be represented using atomic orbitals whose site-to-site phase progression is e^{ika}. Let us now obtain an expression for $E(k)$ in terms of matrix elements whose values can be parameterized in a way that yields a neat form for $E(k)$. Start by writing $E(k)$ as the expectation value of H over $\psi_k(x)$:

$$E(k) = \langle \psi_k(x) | H | \psi_k(x) \rangle \tag{4.55}$$

$$= N^{-1} \sum_{i,i'} e^{ik(x_i - x_{i'})} \langle \phi_\nu(x - x_{i'}) | H | \phi_\nu(x - x_i) \rangle . \tag{4.56}$$

Without carrying out any calculation, we know that, because of the fact that $E(k)$ is real, a phase factor cannot appear in the final result. The $e^{ik(x_i - x_{i'})}$ term will not appear alone following summation.

The double sum can be reduced to a single sum. To see how this works, consider Fig. 10, which indicates the strategy. Focusing on the orbital with $x_{i'} = 0$ [$\phi_\nu(x - x_{i'}) = \phi_\nu(x)$], its interactions with adjacent orbitals are $\langle \phi_\nu(x) | H | \phi_\nu(x \mp a) \rangle e^{\pm ika}$, and so on for its interactions with orbitals centered at $\pm 2a, \pm 3a \ldots$.

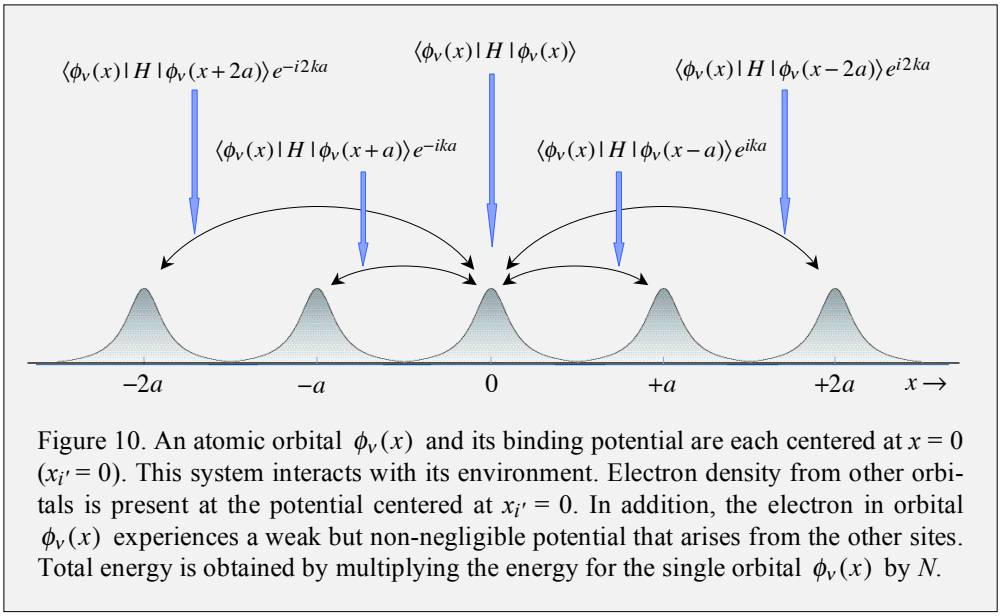


Figure 10. An atomic orbital $\phi_\nu(x)$ and its binding potential are each centered at $x = 0$ ($x_{i'} = 0$). This system interacts with its environment. Electron density from other orbitals is present at the potential centered at $x_{i'} = 0$. In addition, the electron in orbital $\phi_\nu(x)$ experiences a weak but non-negligible potential that arises from the other sites. Total energy is obtained by multiplying the energy for the single orbital $\phi_\nu(x)$ by N.

There is nothing special about the choice $x_{i'} = 0$, so the total energy is obtained by multiplying the result obtained at $x_{i'} = 0$ by N. That is, summing over i' is equivalent to multiplication by N. This cancels the factor of N^{-1} in eqn (4.56) and yields

$$E(k) = \sum_i e^{ikx_i} \langle \phi_\nu(x) | H | \phi_\nu(x - x_i) \rangle \tag{4.57}$$

$$= \langle \phi_\nu(x) | H | \phi_\nu(x) \rangle + \sum_{i \neq 0} e^{ikx_i} \langle \phi_\nu(x) | H | \phi_\nu(x - x_i) \rangle . \tag{4.58}$$

Evaluation of the first term in eqn (4.58) is facilitated by splitting H into $H_0 + V'(x)$, where H_0 is local to the site under consideration. It uses just the potential due to the ion present at $x = 0$, whereas $V'(x)$ takes into account the extended ranges of the potentials centered at the different sites.

$$H = H_0 + V'(x) \tag{4.59}$$

$$= \frac{p^2}{2m} + V_0(x) + V'(x) \tag{4.60}$$

Referring to Fig. 11(a), the potential due to the ion that is present at the site under consideration is $V_0(x)$. Thus, the atomic Hamiltonian for the electron is H_0. The term $V'(x)$ in Fig. 11(b) is the potential at the site under consideration due to the other sites. For the "other sites," we shall use just the immediately adjacent ones because in the tight binding approximation the potentials act rather locally in the regions around their respective ion cores. Consequently, sites farther away than a single displacement a are neglected. With H given by eqn (4.60), the first term on the right hand side of eqn (4.58) becomes

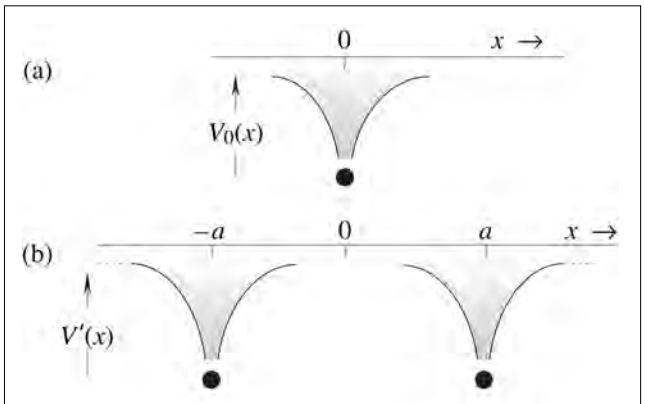


Figure 11. (a) An electron is bound to an ion at $x = 0$. (b) Potentials due to ions at $x = \pm a$ are weak but non-negligible in the vicinity of $x = 0$.

$$\langle \phi_v \mid H \mid \phi_v \rangle = \langle \phi_v \mid H_0 \mid \phi_v \rangle + \langle \phi_v \mid V' \mid \phi_v \rangle \tag{4.61}$$

$$= E_v - \beta \tag{4.62}$$

where E_v is the energy of the atomic orbital ϕ_v. The convention is that $E = 0$ corresponds to the electron removed completely from the atomic core. Thus, all of the atom's bound levels have negative energies. The term $-\beta$ is the average value of $V'(x)$. Because $V'(x)$ is negative, β is positive.

Let us now examine the second term on the right hand side of eqn (4.58). As mentioned earlier, the interaction terms are non-negligible only for adjacent sites in the tight binding model, so the sum reduces to two terms whose matrix elements are

$$\left\langle \phi_v(x) \right| H_0 + V'(x) \left| \phi_v(x \pm a) \right\rangle. \tag{4.63}$$

The H_0 contribution is equal to $E_v \langle \phi_v(x) | \phi_v(x \pm a) \rangle$, which we label $-\gamma$, where γ is positive. Though the magnitude of this term is small, it is not negligible. It is the same order-of-magnitude as β. On the other hand, the $V'(x)$ contribution in eqn (4.63) is sufficiently small that it can be neglected in the qualitative model under consideration, because it involves the product of two small factors: the overlap of $\phi_v(x)$ and $\phi_v(x \pm a)$, and the small magnitude of $V'(x)$ in the vicinity of the site under consideration. Equation (4.58) can thus be written

$$E(k) = E_v - \beta - \gamma \left(e^{ika} + e^{-ika} \right) \tag{4.64}$$

$$= E_v - \beta - 2\gamma \cos ka \, . \tag{4.65}$$

Using $\cos ka = 1 - 2\sin^2(ka/2)$, and defining $E_0 = E_v - \beta - 2\gamma$, eqn (4.65) becomes

$$\boxed{E(k) = E_0 + 4\gamma \sin^2 \left(\tfrac{1}{2} ka \right).} \tag{4.66}$$

Figure 12 is a plot of this dispersion relation. The width of the band is 4γ, which is small. As the overlap integral: $\langle \phi_v(x) | \phi_v(x \pm a) \rangle$, diminishes in magnitude, so does the width of the band. As discussed earlier, the curvature of $E(k)$ can be used to define an effective mass m^*. Because the energy band is narrow in the tight binding limit, energy depends only weakly on k. Therefore, near the bottom of the dispersion curve ($k \sim 0$) curvature is small and the effective mass is larger than that of a free electron. This is a manifestation of the fact that it is hard for the electron to go anywhere when it is bound tightly to a site. In other words, it appears to be sluggish.

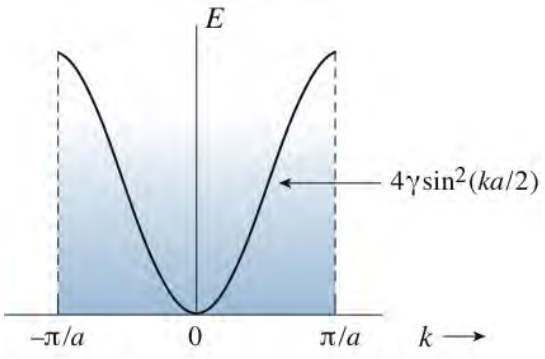


Figure 12. Equation (4.66) is the dispersion relation for the tight binding limit: $E(k) = E_0 + 4\gamma \sin^2(ka/2)$. The energy width is 4γ.

On the other hand, near $k = \pm \pi / a$, where the curvature is negative, the effective mass is negative. It has the same magnitude as at $k = 0$, because the curvature of sine-squared in eqn (4.66) differs only in sign for the stationary points $k = 0$ and $\pm \pi / a$.

An expression for m^* near the stationary points in the dispersion curve is obtained from eqn (4.66) by following the same procedure used in the nearly free electron model. For $k \sim 0$, replacing $\sin(ka/2)$ with $ka/2$ gives

$$\left(\frac{\partial^2 E(k)}{\partial k^2}\right)_0 = 2\gamma a^2 . \tag{4.67}$$

With this set equal to $\hbar^2/m^*$, the effective mass is

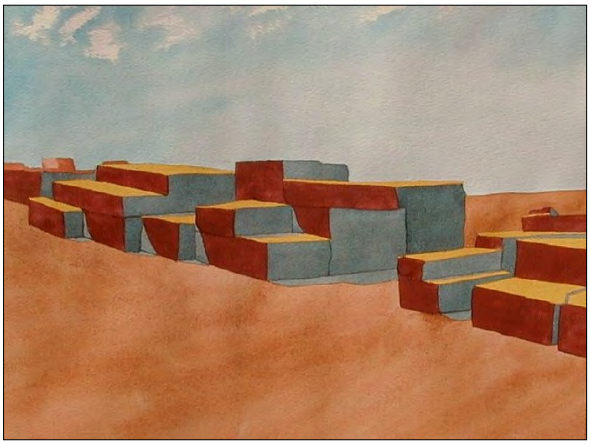
$$\boxed{m^* = \frac{\hbar^2}{2\gamma a^2}}, \tag{4.68}$$

with the understanding that m^* is positive near $k = 0$ and negative near $k = \pi/a$.

This completes our peek into electrons in a 1D periodic lattice, including the limiting cases of nearly free electron and tight binding.

Chapter 5. Simple Molecular Orbital View

We now turn to systems in which the molecular character of site-to-site interactions is central. These are close in spirit to the Hückel model and tight binding limit. The Bloch theorem is exact, so it can be applied to any number of evenly spaced sites, from small to infinite. The goal of this section is to illustrate consequences of the Bloch theorem in systems (small and large) where the molecular character of the bonding between adjacent sites is central. This section is short, offering but a glimpse into a fascinating realm. It draws from the book by Hoffmann [13].



Building Blocks
Robert Wittig

To begin, $N = 3$ is revisited from a perspective that shows how orbital symmetry can destabilize the 3-fold symmetric configuration, resulting in Jahn-Teller distortion. Following this, we will consider the case of 6-fold potentials, linear chains of *s* and *p*-orbitals, and 2D lattices of *s* and *p*-orbitals.

Threefold Symmetry

Figure 1 indicates three symmetrically placed identical orbitals (χ_i) and their respective Bloch phase factors: e^0 (= 1), e^{ika}, and e^{i2ka}. Complex molecular orbitals are constructed using different *k* values. Figures 1(b) and (c) are for *s*-orbitals. The $k = 0$ orbital in (b) is totally symmetric (*a*-type) with no nodes; it has the lowest energy. The next *k* values are $\pm 2\pi/3a$, which correspond to a wavelength of $3a$. Putting these in the Bloch phases yields the following combination of atomic orbitals:

$$\phi_{2\pi/3a} = \chi_1 + \chi_2 e^{i2\pi/3} + \chi_3 e^{i4\pi/3}. \tag{5.1}$$

$$\phi_{-2\pi/3a} = \chi_1 + \chi_2 e^{-i2\pi/3} + \chi_3 e^{-i4\pi/3}. \tag{5.2}$$

Being degenerate (*e*-type), any combination of these can be used. We shall now construct real orbitals. Subtraction of $\phi_{-2\pi/3a}$ from $\phi_{2\pi/3a}$ yields the upper entry in Fig. 1(c), while addition yields the lower entry. Analysis of the three normal modes of this system yields a similar result. One is totally symmetric (*a*-type) and the other two are degenerate (*e*-type). Being degenerate, the latter two can be combined in such a way that their displacements are commensurate with the nodal structure of the electron orbitals. The most common representation of these vibrations labels the normal mode coordinates

Q_a and Q_b and defines them according to Fig. 2. Nuclear displacements can lower the energy of the system by lessening the curvature of the electron wave function that is imposed by the nodes. This constitutes an orbital instability, as explained below.

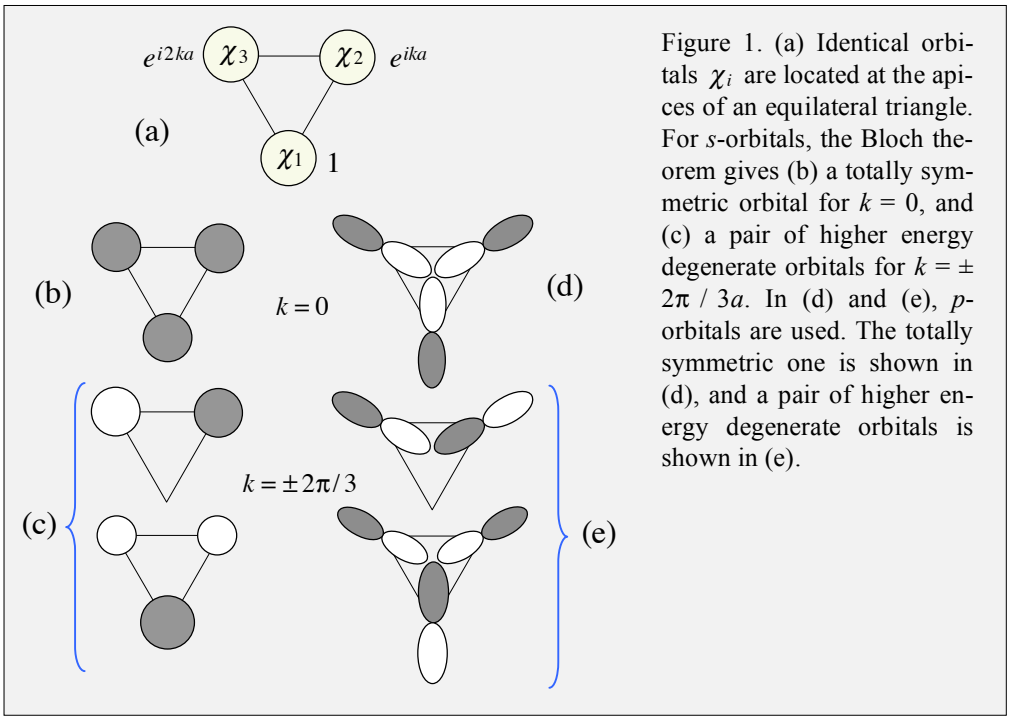


Figure 1. (a) Identical orbitals χ_i are located at the apices of an equilateral triangle. For *s*-orbitals, the Bloch theorem gives (b) a totally symmetric orbital for $k = 0$, and (c) a pair of higher energy degenerate orbitals for $k = \pm 2\pi / 3a$. In (d) and (e), *p*-orbitals are used. The totally symmetric one is shown in (d), and a pair of higher energy degenerate orbitals is shown in (e).

The system indicated in Fig. 1(a) is generic in the sense that no specific orbital has been mentioned. Suppose that each orbital is hydrogen 1*s*. Two electrons fill the lowest orbital (b) and one enters the degenerate pair of orbitals (c), resulting in a ground state that is doubly degenerate. The ground state of H_3 is clearly unstable, progressing toward $H + H_2$. As mentioned above, the nuclear displacements needed to advance the system toward $H + H_2$ are described with the two degenerate modes, which can be made commensurate with the nodal structure.

In the case of H_3, orbital instability of the symmetric arrangement is drastic. The ground state degeneracy is removed spontaneously by the system going to $H + H_2$. Were one to study the exchange reaction H' + H"H"' → H'H" + H"', degeneracy on the H_3 potential surface might play a role, for example, insofar as geometric phase. This orbital instability of H_3 is an example of the Jahn-Teller effect. Usually, orbital instabilities are not so dramatic. They result in a lowering of the symmetry but without chemical change.

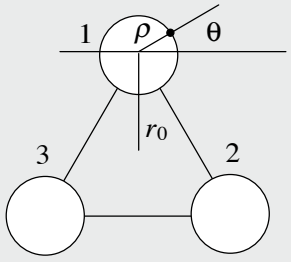


Figure 2. Nuclei circulate on small circular orbits of radii ρ. Though they remain localized at their respective equilibrium sites, angular momentum arises through this synchronous motion, which is in effect a distortion wave circulating on the periphery.

If we now try $k = \pm 4\pi/3a$, this yields

$$\phi_{4\pi/3a} = \chi_1 + \chi_2 e^{i4\pi/3} + \chi_3 e^{i8\pi/3} \tag{5.3}$$

$$\phi_{-4\pi/3a} = \chi_1 + \chi_2 e^{-i4\pi/3} + \chi_3 e^{-i8\pi/3}. \tag{5.4}$$

Higher k values follow the same prescription. If one adds and subtracts eqns (5.3) and (5.4), the orbitals in Fig. 1(c) are again obtained. In other words, $\phi_{4\pi/3a}$ is equal to $\phi_{-2\pi/3a}$ and $\phi_{-4\pi/3a}$ is equal to $\phi_{2\pi/3a}$. As expected, these are not new orbitals, as only 3 independent orbitals are possible.

In (d) and (e), *p*-orbitals are used. Nothing much changes, as the Bloch theorem gives the same site-to-site phase progression. Namely, the orbitals shown in (d) and (e) are a totally symmetric one and a doubly degenerate pair, respectively, and the system undergoes a Jahn-Teller distortion. The general result does not depend on details of the atomic orbitals, only the 3-fold symmetry.

Sixfold Symmetry

Let us now apply the Bloch theorem to the 6-fold system indicated in Fig. 3. The circles with dots at their centers can be taken to be either *s*-orbitals or p_z-orbitals coming out of the paper (all of one or all of the other), the latter being appropriate for benzene. For the degenerate pairs, molecular orbitals will be constructed to be real. The small white and dark grey circles indicate position and relative phase. Their sizes are not scaled to probability densities. For $k = 0$, the molecular orbital is totally symmetric. For $k = \pi/3a$ the wavelength ($\lambda = 2\pi/k$) is equal to $6a$ so there are two nodes. For $k = 2\pi/3a$ the wavelength is $3a$ so there are 4 nodes. For $k = \pi/a$, which is the edge of the Brillouin zone, there are 6 nodes. The procedure for obtaining these results is given in the box below. It is easily extended to other symmetrically spaced orbitals on a ring.

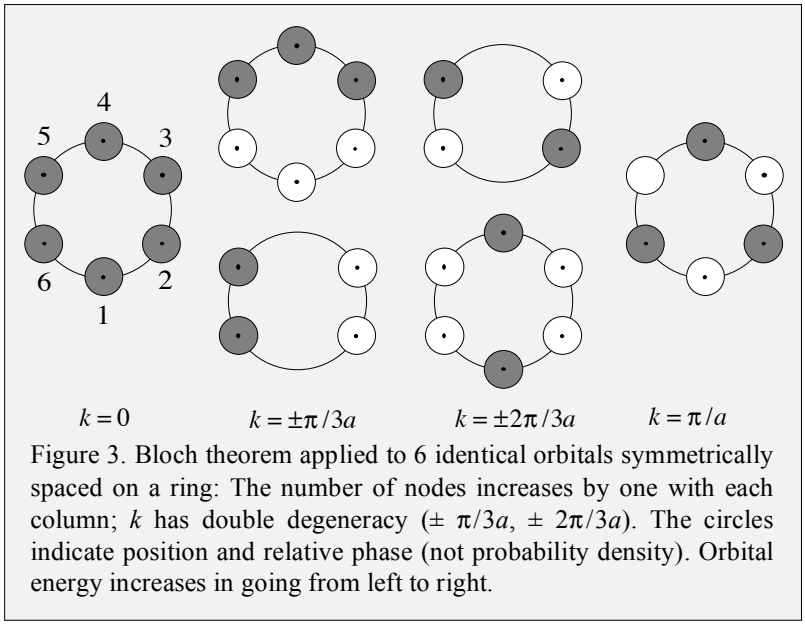


Figure 3. Bloch theorem applied to 6 identical orbitals symmetrically spaced on a ring: The number of nodes increases by one with each column; k has double degeneracy ($\pm\ \pi/3a$, $\pm\ 2\pi/3a$). The circles indicate position and relative phase (not probability density). Orbital energy increases in going from left to right.

For the case $k = \pi/a$, which is the edge of the Brillouin zone, λ matches the periodicity of the potential, namely, $\lambda = 2a$. Recall the square-well band gaps examined in Chapter 1. There we saw that the upper and lower states at the band gap correspond to probability density peaking above the high and low parts, respectively, of the 6-fold potential. The case shown in Fig. 3 ($k = \pi/a$) is analogous.

Bloch Theorem Applied to Six Orbitals

Applying Bloch's theorem to a particle in a sixfold periodic potential on a ring yields

$$\psi = \chi_1 + \chi_2 e^{ika} + \chi_3 e^{i2ka} + \chi_4 e^{i3ka} + \chi_5 e^{i4ka} + \chi_6 e^{i5ka}.$$

Allowed values of k up to the edge of the first Brillouin zone gives six orbitals:

$k = 0$ $\quad \chi_1 + \chi_2 + \chi_3 + \chi_4 + \chi_5 + \chi_6$

$k = \pi/3a$ $\quad \chi_1 + \chi_2 e^{i\pi/3} + \chi_3 e^{i2\pi/3} - \chi_4 + \chi_5 e^{i4\pi/3} + \chi_6 e^{i5\pi/3}$

$k = -\pi/3a$ $\quad \chi_1 + \chi_2 e^{-i\pi/3} + \chi_3 e^{-i2\pi/3} - \chi_4 + \chi_5 e^{-i4\pi/3} + \chi_6 e^{-i5\pi/3}$

$k = 2\pi/3a$ $\quad \chi_1 + \chi_2 e^{i2\pi/3} + \chi_3 e^{i4\pi/3} + \chi_4 + \chi_5 e^{i2\pi/3} + \chi_6 e^{i4\pi/3}$

$k = -2\pi/3a$ $\quad \chi_1 + \chi_2 e^{-i2\pi/3} + \chi_3 e^{-i4\pi/3} + \chi_4 + \chi_5 e^{-i2\pi/3} + \chi_6 e^{-i4\pi/3}$

$k = \pi/a$ $\quad \chi_1 - \chi_2 + \chi_3 - \chi_4 + \chi_5 - \chi_6$

Real molecular orbitals are obtained by taking combinations that differ only in the sign of k. For $k = \pm \pi/3a$, adding the above Bloch orbitals yields

$$2\chi_1 + \chi_2\left(e^{i\pi/3} + e^{-i\pi/3}\right) + \chi_3\left(e^{i2\pi/3} + e^{-i2\pi/3}\right) - 2\chi_4 + \chi_5\left(e^{i4\pi/3} + e^{-i4\pi/3}\right) + \chi_6\left(e^{i5\pi/3} + e^{-i5\pi/3}\right)$$

$$= 2\left(\chi_1 + \chi_2 \cos 60 + \chi_3 \cos 120 - \chi_4 + \chi_5 \cos 240 + \chi_6 \cos 300\right)$$

$$= 2\left(\chi_1 + \tfrac{1}{2}\chi_2 - \tfrac{1}{2}\chi_3 - \chi_4 - \tfrac{1}{2}\chi_5 + \tfrac{1}{2}\chi_6\right)$$

This orbital is shown in Fig. 3 ($k = \pm\pi/3a$, upper); the lower entry is obtained by subtracting. Some algebra yields $i\sqrt{3}\left(\chi_2 + \chi_3 - \chi_5 - \chi_6\right)$. The result for the other degenerate pair follows along the same lines. Cases of zero and six nodes are written by inspection.

For a linear chain of atoms, how energy varies with k depends on the nature of the orbitals. Figure 4 shows s-orbitals and in-plane p-orbitals. The former are bonding for $k = 0$ and antibonding for $k = \pi/a$. The latter are the opposite: antibonding for $k = 0$ and bonding for $k = \pi/a$. This means that if E versus k is plotted, the curve will increase with k for the s-orbitals, but decrease with k for the p-orbitals. The same considerations apply to d-orbitals. It would be necessary to examine the bonding characteristics of the different members of the d-block to see the bonding, antibonding, and other propensities and how they vary with k.

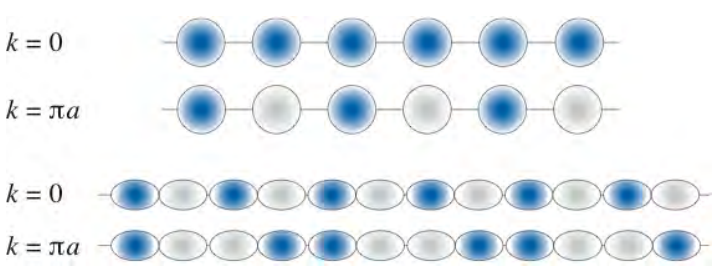


Figure 4. Linear chains showing $k = 0$ and π/a. Whether the energy increases or decreases with k depends on the orbitals. For s-orbitals, $k = 0$ is bonding and $k = \pi/a$ is antibonding. For the p-orbitals shown, $k = 0$ is antibonding and π/a is bonding.

Two Dimensions

Let us now extend the model to 2D. In Fig. 5 a lattice is introduced in which separations between adjacent sites are the same in the $\hat{x}$ and $\hat{y}$ directions. A 6×6 segment of the infinite 2D crystal suffices to illustrate the vector nature of $\vec{k} = k_x \hat{x} + k_y \hat{y}$. The situation is symmetrical in $\vec{k}$ space. The first Brillouin zone is contained within a square in $\vec{k}$ space whose sides are of length $2\pi/a$. All of the information can be given in a quadrant whose sides have length π/a. Entries (a)–(g) show atomic s-orbitals for a few values of k_x and k_y, while entries (h)–(k) use p_x and p_y.

Consider first the s-orbitals. When $k_x = k_y = 0$, the crystal orbital is totally symmetric, as shown in (a). When $k_x = \pi/2a$ and $k_y = 0$, the Bloch theorem yields, for each value of y, the following linear combinations of atomic orbitals:

$$\chi_1 + i\chi_2 - \chi_3 - i\chi_4 + \chi_5 + i\chi_6 \tag{5.5}$$

$$\chi_1 - i\chi_2 - \chi_3 + i\chi_4 + \chi_5 - i\chi_6 \tag{5.6}$$

Adding and subtracting gives, to within overall multiplicative factors,

$$\chi_1 - \chi_3 + \chi_5 \quad \text{and} \quad \chi_2 - \chi_4 + \chi_6 \tag{5.7}$$

These degenerate orbitals, which extend throughout the crystal, differ only in that they occupy adjacent sites. The first is shown in Fig. 4(b); the second is like it, but displaced horizontally by a. Likewise, when $k_x = 0$ and $k_y = \pi/2a$, application of the Bloch theorem results in two degenerate orbitals, one of which is shown in (c).

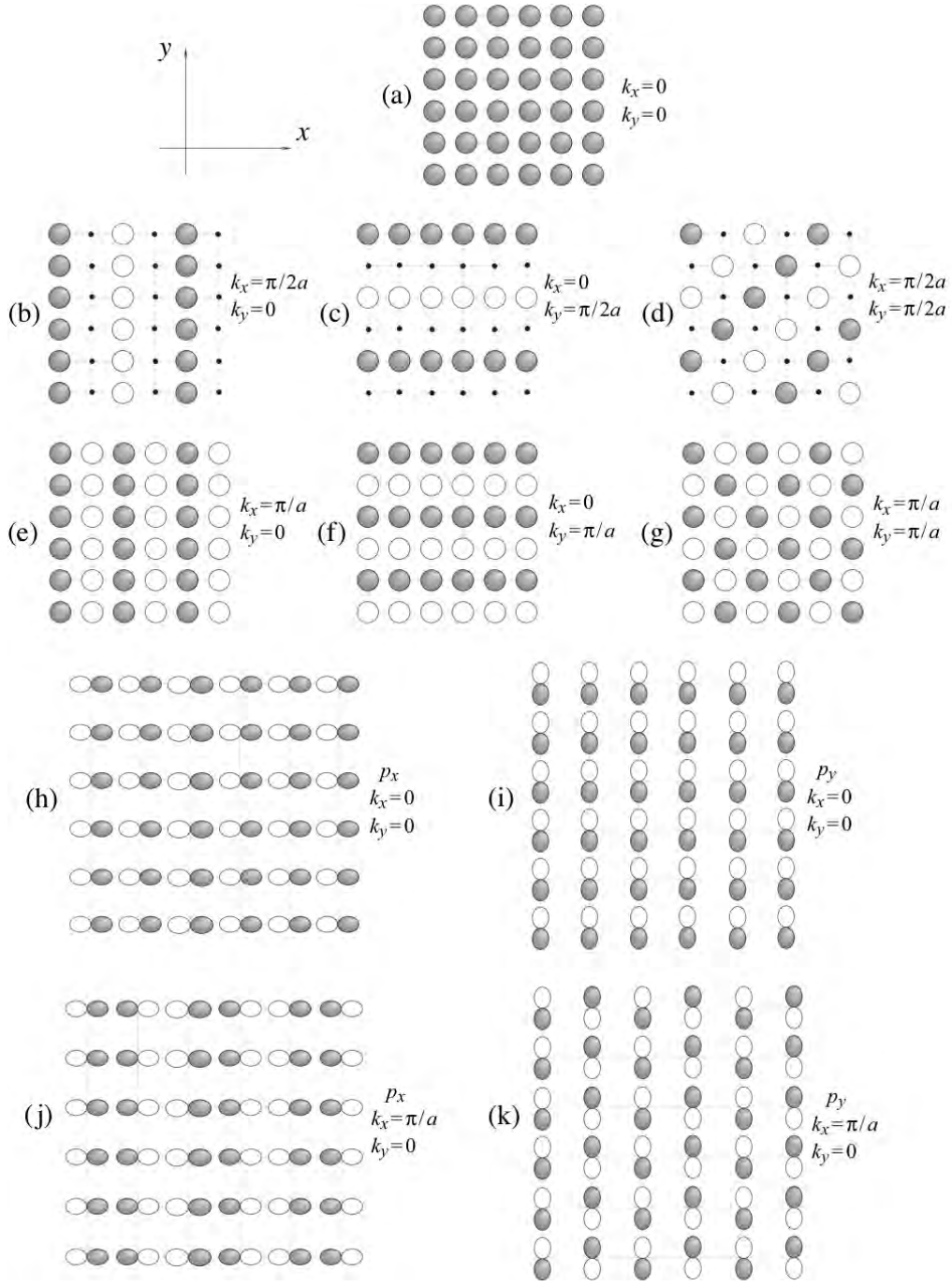


Figure 4. Square 6×6 portions of infinite 2D lattices are indicated. Entries (a) – (g) show *s*-orbitals for values of k_x and k_y, whereas (h) – (k) show *p*-orbitals.

In (d), lines of constant phase lie at 45° with respect to the *x*-axis. The gradient of the phase has its largest magnitude in a direction that is perpendicular to these lines of constant phase. There are four possible crystal orbitals for $k_x = \pm \pi/2a$ and $k_y = \pm \pi/2a$. They are obtained from (d) by exchanging the filled and unfilled circles and/or rotating the matrix of circles by 90° relative to the lattice. An important point is that the direction of $\vec{k}$ defines the nodal structure.

Entries (e) – (g) differ from (b) – (d) only in having $\vec{k}$ twice as large. They lie at the edge of the Brillouin zone, which is a square in $\vec{k}$-space . Entry (g), which lies at the corner of the square in $\vec{k}$-space , is the highest energy possible for the given set of *s*-orbitals. To get to higher energies, it is necessary to introduce excited states of the atoms.

Part A. Bibliography and References

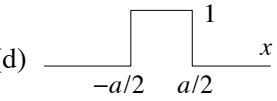
1. N. W. Ashcroft and M. D. Mermin, Solid State Physics (Brooks-Cole, Pacific Grove, CA, 1976).
2. J. S. Blackstone, Solid State Physics, Second Edition (Cambridge University Press, Cambridge, 1998).
3. J. R. Hook and H. E. Hall, Solid State Physics, Second Edition (Wiley, New York, 2004).
4. H. Ibach and H. Lüth, Solid State Physics (Springer Verlag, Berlin, 1993).
5. C. Kittel, Introduction to Solid State Physics, 8th Ed. (Wiley, New York, 2005).
6. W. A. Harrison, Solid State Theory (Dover, New York, 1980).
7. M. A. Omar, Elementary Solid State Physics (Addison Wesley, New York, 1975).
8. O. Madelung, Introduction to Solid-State Theory (Springer, New York, 1996).
9. J. M. Ziman, Electrons and Phonons (Oxford University Press, Oxford, 2001).
10. J. M. Ziman, Theory of Solids (Cambridge University Press, Cambridge, UK, 1972).
11. M. Born and K. Huang, Dynamical Theory of Crystal Lattices (Clarendon Press, Oxford, 1954, reprinted 2002).
12. C. M. Wolfe, N. Holonyak, Jr., and G. E. Stillman, Physical Properties of Semiconductors (Prentice Hall, Englewood, New Jersey, 1989).
13. R. Hoffmann, Solids and Surfaces: A Chemist's View of Bonding and Extended Structures (Wiley VCH, New York, 1988).
14. L. Brillouin, Wave Propagation in Periodic Structures (Dover, Mineola, New York, 2003), reprinting of work published in 1946.
15. A. F. J. Levi, Applied Quantum Mechanics (Cambridge University Press, Cambridge, England, 2003).
16. R. Gilmore, Elementary Quantum Mechanics in One Dimension (Johns Hopkins University Press, Baltimore MD, 2004).
17. G. B. Arfken and H. J. Weber, Mathematical Methods for Physicists, Seventh Edition (Academic Press, New York, 2012).
18. B. Kusse and E. Westwig, Mathematical Physics: Applied Mathematics for Scientists and Engineers (Wiley, New York, 2006).
19. J. Matthews and R. L. Walker, Mathematical Methods of Physics, Second Edition (Addison-Wesley, New York, 1970).
20. J. W. Brown and R. V. Churchill, Complex Variables and Applications, Ninth Edition (McGraw-Hill, New York, 2013).

Part A. Exercises and Projects

1. Chapter 1: Consider a POR with potential $V_0 \sin 24\phi$. Each of the 24 sites contributes 2 electrons. The ring has a radius of 40 Å, $V_0 = 1000$ cm^{-1}, and $T = 300$ K. Discuss whether the 2×2 matrix approach can be used, or is it necessary to diagonalize a large matrix. Make a plot showing the orbital energies and the extent to which they are occupied. In other words, plot $n(E)$ versus E, where $n(E)$ is the number of electrons having energy E. Keep in mind that electrons obey Fermi-Dirac statistics. How does the electrical conductivity vary with temperature? What effect would impurities that donate electrons have on the conduction of this system?

2. Chapter 1: Facility with Fourier transformation is an essential tool. Evaluate the Fourier transforms (and make sketches of the results) for the following functions:

   (a) $e^{-\gamma t}$ $(t \geq 0)$; (b) $e^{-(x/x_0)^2}$; (c) $\frac{1}{\pi} \frac{\gamma}{(\omega - \omega_0)^2 + \gamma^2}$ (d)

3. Chapter 1: For a discrete lattice, show by explicit calculation that: $L^{-1} \int_0^L dx\, e^{ikx} e^{ik'x}$ vanishes when $k + k' \neq 0$. What is the value when $k + k' = 0$?

4. Chapter 1: Make plots of the 60 lowest eigenvalues for $V(\phi) = V_0 \cos^4 12\phi$ for two limiting cases. Namely, choose parameter values such that the limits of weak and strong binding are examined. Discuss the results. Keep in mind that for strong binding a large POR basis is required.

5. Chapter 1: Consider $V(\phi)$ with 6 repeating segments. You are allowed to choose a maximum of 3 Fourier coefficients in order to adjust the shape of the potential. Construct $V(\phi)$ such that it has two bands with $E < 0$, and the rest have $E > 0$. Examine the bands with $E < 0$. Plot $|\psi|^2$ for each of the 6 eigenstates in each of the bands. Discuss the results.

6. Chapter 1: Figures 12(b) and (c) display undulations in which the amplitudes vary from site to site. Use the Bloch theorem to derive the behavior seen in (b) by finding the appropriate orbitals. Make predictions for the probability densities for $k = \pi/6a$ and $\pi/3a$.

7. Chapter 1: Examine two approaches to the nearly free electron limit of the 24 site POR model: (i) matrix diagonalization and (ii) 2×2 matrices with pairs that satisfy $|m - m'| = N$. Make plots like Fig. 7(c). Vary the strength of the potential until the approaches differ and discuss why they differ. Are the eigenvalues a good reflection of free electron behavior? Would the eigenvectors be better?

8. Chapter 1: Devise a potential $V(\phi)$ for 12 sites that has 4 bands with $E < 0$. The choice of Fourier coefficients in the expansion of $V(\phi)$ is yours. Plot the eigenvalues

in a manner similar to Fig. 11, except here there will be more low-lying bands. Now the fun begins. Plot $|\psi|^2$ versus ϕ for 4 eigenstates from each of the two lowest bands. You choose the eigenstates. Discuss the results.

6. This exercise concerns density of states. Write the 1D *k*-space density for the POR and convert this to $\rho(E)$ using $E = 2\beta \cos ka$. Does this form of $\rho(E)$ make sense? What is obtained if $\rho(E)$ is integrated over the first Brillouin zone? Extend the density in *k*-space to 2D, then 3D. There is nothing special about any direction, so you should integrate over angles in *k*-space. Make the bold assumption that $E = 2\beta \cos ka$ in the first Brillouin zone of the 3D system, and give an expression for $\rho(E)$.

7. Explain why $\det T = 1$ is justified in the transfer matrix method.

8. Consider the Hückel model applied to a system consisting of 8 sp^2 carbon atoms. They can be arranged in a ring or linear chain. Discuss the molecular orbitals and stabilities. If instead of 8 we have 1000 sp^2 carbon atoms, will this be a conductor or an insulator? Explain.

9. Consider electrons in a 1D lattice of length *L*. Show that $\rho(E)$ can be expressed as: $\rho(E) = L / hv_g$, where *h* is Planck's constant and v_g is the group velocity. Discuss the behavior of $\rho(E)$ at the band edges and how it varies between $E = 0$ and the first band edge. How does this differ for the nearly free electron and tight binding limits?

10. Electrons are in a sufficiently weak potential: $V(x) = A\cos k_0 x + B\cos 2k_0 x$ (*A* and *B* are real) that the nearly free electron approximation applies. What are the effective masses at the first and second band gaps? Are there higher band gaps? If so, discuss their character. Make a plot of v_g versus *k*.

11. Can bands overlap with a 1D periodic potential. If not, why?

12. Consider a 1D periodic lattice made up of delta function potentials $V_0 \delta(x - na)$, where *n* is an integer. Make plots of $E(k)$ versus *k* for several values of the strength parameter V_0. Discuss these dispersion relations. For the nearly free electron, what are the sizes of the energy gaps?

13. Consider the Bloch wave functions: $\psi_{k,n} = u_{k,n}(x)e^{ikx}$, for the case $u_{k,n}(x) \propto (x - ia)\exp(-(x - ia)^2)$, where *i* is an integer. Clearly, $u_{k,n}(x)$ goes to zero at $x = ia$. Does this mean the flux is equal to zero? Explain.

14. Is it possible to have a periodic potential (of appreciable magnitude) that does not have a band gap at $k = \pm\pi/a$?

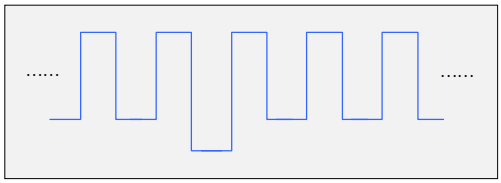
15. Consider $\rho(E)$ in the nearly free electron and tight binding limits. In which of these limits is $\rho(E)$ higher? Discuss.

16. Consider the function: $N^{-1/2}\sum_s e^{ika}u_k(x-x_s)$. Is it a legitimate Bloch function?

17. You have 1D PIB and POR densities of states. Now increase the dimension. For the PIB this is easy conceptually. Give the densities of states for 2D and 3D. Now go to the ring. The ring is a 2D object but the radius is fixed so motion is 1D. What is the density of states for a particle on the surface of a sphere?

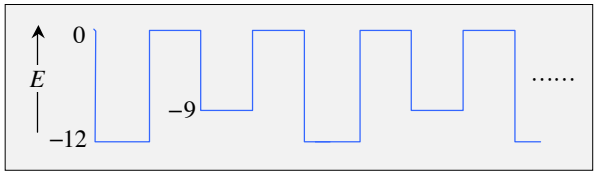
Projects

1. This is about an impurity in a periodic structure. Set up a POR with 12 repeating elements; a rectangular potential will suffice. Now make one of the wells more binding, as indicated in the figure. The idea is to explore the energy levels of this system. Specifically, will the different site create new levels? How will the different site affect the band structure of the rest of the lattice? To pursue these issues, use any method you like to solve for the energies, and find energies and bands for levels below the barrier height and also above it. This project is open ended, so you will need to carefully discuss what is going on.



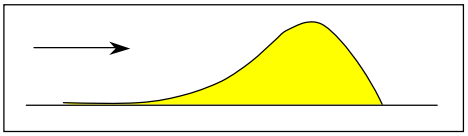
2. This is about the periodic potentials: POR in the tight binding limit; and Hückel. Set up and solve for a POR with a 12-fold potential in the tight binding limit. The strength of the potential is not important as long as it yields the tight binding limit for the lowest band. List eigenvalues and plot wave functions. Now go to Hückel theory. Make a circle of 12 atoms and calculate the energies and wave functions (values of α and β from the literature). Repeat for a conjugated polyene of 12 atoms. Compare the wave functions for the two models. Discuss similarities and differences. Are there significant differences if 11 sites are used instead of 12? Discuss odd versus even.

3. The strategy behind this project is to use a simple 1D model to learn something about an alloy in which alternate binding sites differ. Start by setting up a POR model for the potential shown. Use 12 repeating elements. Well widths are 4 Å and spaces between wells are also 4 Å. Energies are in eV and the particle is an electron. This can be solved using the transfer matrix method. Find the energies of all states with $E < 0$, and some of the states with $E > 0$. Repeat with a more rounded potential. Check with me after the first part and we will decide on how to proceed with this second part.



4. This project is about tunneling, namely, dependence on mass, barrier strength, barrier shape, and incident energy. To begin, devise a shape for the barrier, for example, like

the one indicated in the figure. For masses, use electron, proton, and 20 amu. Set up the transfer matrix program and make a few runs for an electron. Use a barrier height of 1 eV, and plot transmission probability versus incident energy. Show me the results and we will go from there. The idea is to gain an intuitive grasp of the relationship between the molecular parameters and the efficacy of the tunneling mechanism, which is important in a number of systems.

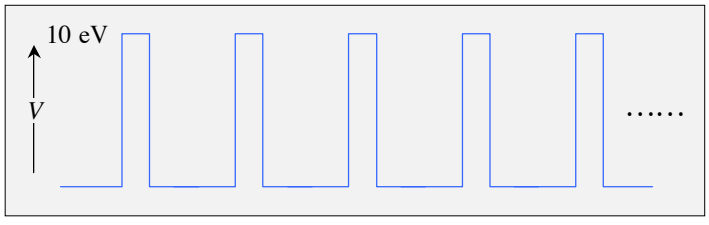


5. Consider the system shown in the figure. Assume m and m' are 17 and 1 amu, respectively, and $k >> k'$, making this system vaguely reminiscent of water. The large spring constant k and small mass m' give rise to a high frequency vibration (~ OH stretch). Assume a frequency (energy) of 3600 cm^{-1} for the isolated vibration of m' relative to the left wall, and find k. The other couplings have spring constants $k' = k/10$, resulting in significantly lower frequencies. The rightmost m is a friction plate. The force it experiences is equal to a constant times the velocity ($b\dot{x}$). Write the equation of motion. Use $x_n(t) = x_n e^{-i\omega t}$ to set up an eigenvalue problem in the x_n basis, and find the eigenvalues. Pick a value of b that yields a modest exponential decay rate (for example, 10^{12} s^{-1}) for the "mode" that has the highest frequency. Vary b and k' to see how this works.

6. *Vibrational Density of States*: Consider the vibrational densities of states for the triatomic molecules H_2O and NO_2. Start with H_2O, whose normal modes are asymmetric stretch, symmetric stretch, and bend. Use a direct count method to find the vibrational state densities for each symmetry species for energies up to 1 eV above dissociation threshold. Start by assuming harmonic frequencies; then add the first diagonal anharmonicity correction. With these two results in hand, obtain the total (all symmetry species) vibrational density of states by using convolution. Repeat for NO_2, but including the effect of the intersection of potential surfaces.

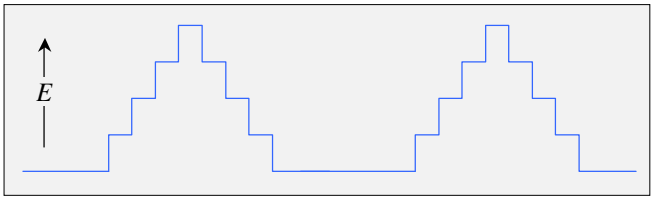
7. Consider a graphene sheet comprised of 40 carbon atoms. Use the Hückel model to obtain the energies of this system (get values of α and β from the literature). Discuss the electrical conductivity and how it varies with the shape of the network. Is it anisotropic? Now add a second graphene sheet, parallel to the first but displaced. The sheets are weakly bound to one another. Construct and diagonalize the corresponding Hamiltonian matrix. Discuss how energies and electrical conductivity are affected?

8. This project is about transmission of an electron through the region indicated in the sketch, specifically, at energies near that of the barrier. Use transfer matrices. Parameters: barrier height 10 eV; barrier width 4 Å; distance between

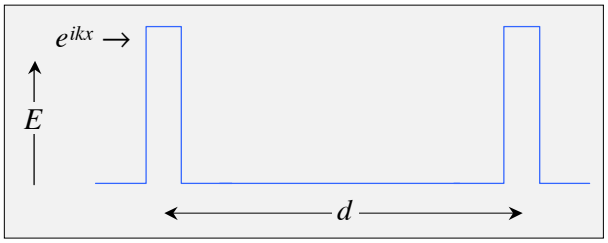


barriers 12 Å; there are 24 repeating elements. For $E < 10$ eV, find the transmission resonance that has the highest energy. Make a careful plot of transmission probability versus E. Now consider $E > 10$ eV. Find the lowest resonance (above 10 eV) and plot transmission efficiency versus E. Discuss the results.

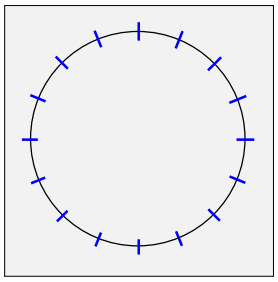
9. This project is about transmission at energies above and below barriers. Use transfer matrices to represent the two barriers indicated in the figure. Do not worry about their precise shape. Make them look more-or-less Gaussian and approximate them with 7 rectangular regions. Make the distance between the barriers several times their width. Calculate transmission probability versus energy. Adjust parameters such that the transmission probability for a single barrier is 0.5 at an energy that is half the barrier height. Extend the result to 8 barriers.



10. This is about resonances (quasibound states). Consider the transmission of a particle wave of variable energy (incident from the left) through a pair of barriers. Make the barrier thickness large enough that at half the barrier height the transmission probability for a single barrier is 0.05. Adjust height and width to achieve this. Also, adjust the distance between barriers such that there are several PIB type states between the barriers. Plot overall transmission (through both barriers) versus energy of the incident wave. What you will find is that states look like PIB states except they have widths. Discuss the widths. What is the relationship to Fourier transforms?



11. Consider the circular loop in the figure. Compartments are separated with impenetrable barriers (short straight lines). To begin, put a single spin-up electron in each of the compartments. Assume the electrons move much more rapidly than do the barriers, and that each electron is in the lowest PIB state. Each barrier has mass m. The barriers are not connected to anything: they are free to move on the circular track of circumference 80 Å without friction. Ignore all Coulomb interactions. In other words, electrons only interact with one another via the barriers.



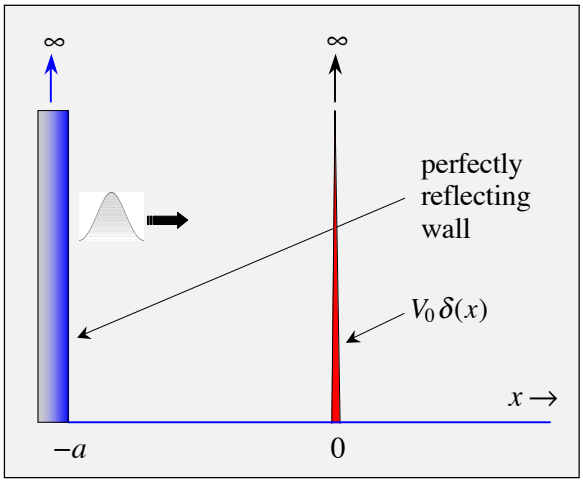
Give the equations of motion for the barriers, retaining just the harmonic term in the potential experienced by the barriers. Write the complete Hamiltonian. Include electron and barrier degrees of freedom. Every other barrier is now removed, in which case there are two spin-up electrons per compartment. Describe qualitatively what

happens. How is the harmonic force constant changed? What is the ground state like? How does the ground state energy compare to that of the 16-compartment system?

12. This project is about resonances. There are two parts: one uses the transfer matrix approach. The other uses the exact solution of a simple model.

Start by reviewing the situation shown in (a), which is assigned as project 10. Explain why transmission can be virtually 100% for energies below the height of two barriers but not with a single such barrier. Next, solve (b) using the transfer matrix approach. Adjust parameters such that there are several bound states in the well. Your result should display transmission resonances in the probability of transmission versus incident energy. Discuss this result.

Now consider the situation on the right. The delta function potential, despite its curious properties, is a useful pedagogical tool. If you are unfamiliar with the use of the delta function in this context, Appendix 3 provides a nice introduction. You should go through it before proceeding.



The situation on the right has resonances whose widths increase with energy, and whose sharpness can be tuned by varying the strength V_0. With the analytical solution in hand, it is child's play to go from isolated to overlapping resonances by increasing E. In addition, a wave packet can be launched in the region $-a \le x \le 0$ to illustrate important effects.

Start with a general expression for the wave function, whose form pays heed to the boundary conditions at $x = -a$ and 0.

$$\psi(x) = A \sin k(x+a) \qquad -a \le x \le 0 \tag{1}$$

$$= A \sin k(x+a) + B \sin kx \qquad x \ge 0\,. \tag{2}$$

The above wave function is continuous at $x = 0$, and its derivative is discontinuous:

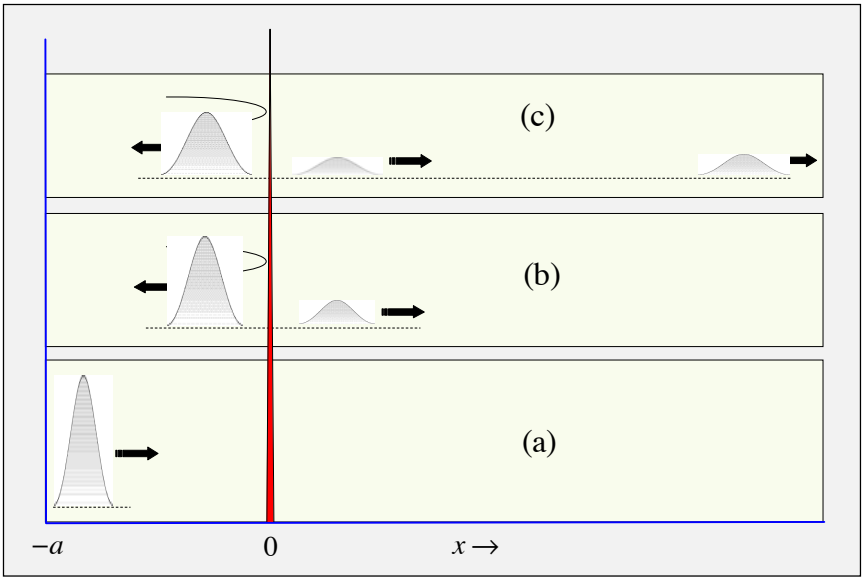
$$\psi'(+\varepsilon) - \psi'(-\varepsilon) = \alpha\psi(0)\,, \tag{3}$$

where $\alpha = 2mV_0/\hbar^2$. The primes denote differentiation with respect to x. Verify that this is true, and that applying it to $\psi(x)$ in eqns (1) and (2) yields $B = (A\alpha/k)\sin ka$, which is put into eqn (2). Normalization is carried out using Dirac delta functions to obtain an expression for A. This tedious exercise yields

$$A = \sqrt{\frac{2/\pi}{1+(\alpha/k)\sin 2ka + (\alpha/k)^2 \sin^2 ka}}\,. \tag{4}$$

Note that the wave functions are not square integrable, as they extend to infinity.

The above expression for A contains most of the important information. Because we have the exact solution, it is safe and instructive to see what happens as parameters are varied over large ranges. At exact resonance ($ka = n\pi$), peaks in $|A|^2$ occur at the same energies as the eigenvalues of a particle-in-an-infinite-well of length a. For a given energy range, increasing the value of $\alpha = 2mV_0/\hbar^2$ causes the resonances to sharpen. Regardless of the value of α, as E increases, the resonances eventually become diffuse, blending into one another. As the energy increases yet further, the regime is reached in which the particles pay little attention to the potential. Make plots that display this information.



Referring to the figure, panel (a) indicates a packet near $x = -a$. No attempt is made at accuracy – the goal is to get the main idea across. The packet moves to the right and when it reaches the potential it is partially reflected and partially transmitted. Panel (b) shows the situation after this bounce. Transmitted and reflected packets spread as they propagate and the packet inside the barrier is partially transmitted with each bounce. A sequence of bursts emerges at $x = 0$, each traveling to the right. Suppose $|R|^2 = 0.8$. After 4 bounces, the probability of finding the particle in the region $-a \le x \le 0$ has dropped to $(0.8)^4$. Of the emerging packets, the first is the

most intense, with each successive burst down by 0.8. Because of the fact that they spread as they propagate, they can blend into one another at long times.

This makes an amusing story, but could something like this happen? The figure below (from the Zewail group) shows a result obtained using fs pulses with what is called the pump-probe technique. The first pulse excites the system, creating a wave packet on the excited potential. The second pulse interrogates the system to see what happens after the packet commences its journey. The molecule is diatomic NaI, so it can be treated as being 1D.

Referring to Fig. 2(a), the repulsive wall of the excited potential is as close to vertical as you are likely to find in a molecule. Likewise, the region between the repulsive wall and the transition region near 6.9 Å is remarkably flat for a molecular potential. Finally, the transition region near 6.9 Å is not a delta function *per se*, but it behaves like one because the system undergoes transitions from the excited state to the ground state in a narrow spatial region. Thus, our toy model is not so far-fetched after all. An experimental result is shown in Fig. 2(b). The pump pulse creates a packet on the left side of the excited potential. It moves to the right, transferring flux to the ground state potential in the region near 6.9 Å. This cycle repeats. When monitoring Na, the signal rises each time the packet reaches the transition region. Alternatively, better resolution is obtained by monitoring the presence of the packet inside the excited potential.

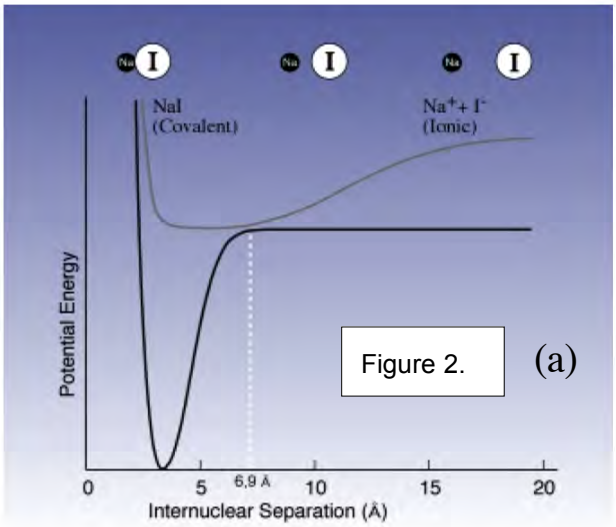
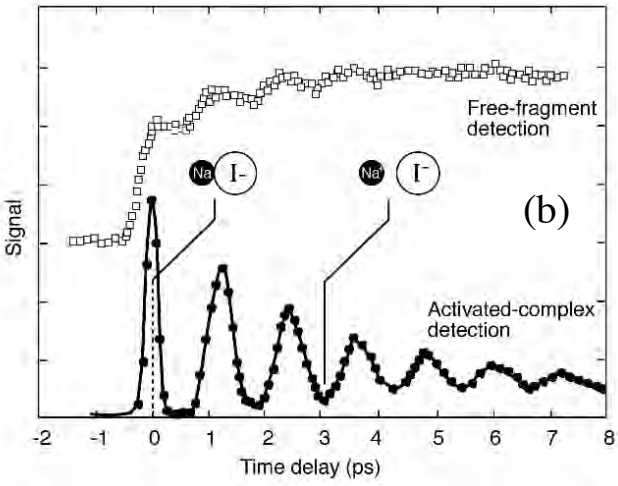


Figure 2.



Part B

Part IV. Appendices