

Lecture I

Contents

I. The variational principle	1
A. The approximate wavefunction	1
B. Variation with respect to coefficients and Lagrange multipliers	2
1. An example : Delta-function potential	3
II. Concepts in periodic systems	5
A. Periodicity and crystals	5
B. Reciprocal space	6
1. An example : Face-centered cubic lattice and its Brillouin zone	7
C. The Brillouin Zone	8
D. Bloch's theorem	8
1. An example : Tight-binding model for atoms on a chain	10
References	11

I. THE VARIATIONAL PRINCIPLE

A. The approximate wavefunction

In quantum mechanics, the main task is to solve the Schrödinger equation.

$$\hat{H}\psi = E\psi \quad (1)$$

The Schrödinger equation is exactly solvable for a very narrow class of systems. In cases where the exact solution cannot be attained, the wavefunction may be approximated by a form that is easier to handle mathematically

$$\phi \approx \psi \quad (2)$$

In most cases, we are interested in the ground state of the system, which we shall denote by ψ_0 yielding the ground state energy, E_0 . The excited states of the system will be denoted by $\{\psi_1, \psi_2, \dots\}$ with corresponding energies $\{E_1, E_2, \dots\}$. In what follows, we will be interested in obtaining an approximation to ψ_0

Unless we are very lucky, the approximate wavefunction ϕ will no longer be an eigenvalue of the Hamiltonian operator, \hat{H} . The quality of the approximation is assessed based on how close the expectation value of \hat{H} for ψ given by

$$\tilde{E} = \frac{\langle \phi | \hat{H} | \phi \rangle}{\langle \phi | \phi \rangle} \quad (3)$$

comes to the actual energy eigenvalue, E_0 . Assuming that the eigenstates of \hat{H} form a complete basis set, we may expand any other wavefunction of the system in terms of them. This observation applies also to our approximation, ϕ . We therefore write down the following expansion

$$|\phi\rangle = \sum_n c_n |\psi_n\rangle \quad (4)$$

where c_n are the expansion coefficients. The eigenstates $\{|\psi_n\rangle\}$ are assumed to be orthonormal. This assumption does not cause us to lose any generality because any complete set of eigenstates may be constructed to be orthonormal. Thus, the following property is satisfied

$$\langle \psi_n | \psi_m \rangle = \delta_{nm} \quad (5)$$

We make the further assumption that the eigenvalues, $\{E_n\}$ are labeled in an increasing order, i.e.

$$E_0 \leq E_1 \leq E_2 \leq \dots \quad (6)$$

Substituting Eq. 4 into Eq. 3, we obtain the following

$$\tilde{E} = \frac{\langle \phi | \hat{H} | \phi \rangle}{\langle \phi | \phi \rangle} = \frac{\sum_{nm} c_n^* c_m \langle \psi_n | \hat{H} | \psi_m \rangle}{\sum_{nm} c_n^* c_m \langle \psi_n | \psi_m \rangle} \quad (7)$$

$$= \frac{\sum_{nm} c_n^* c_m E_n \langle \psi_n | \psi_m \rangle}{\sum_{nm} c_n^* c_m \langle \psi_n | \psi_m \rangle} = \frac{\sum_n |c_n|^2 E_n}{\sum_n |c_n|^2} \quad (8)$$

The last two equalities are a result of the fact that $\{|\psi_n\rangle\}$ are eigenstates of \hat{H} and of orthonormality as stated in Eq. 5. If we now substitute all E_n in Eq. 8 with E_0 , we have

$$\tilde{E} = \frac{\sum_n |c_n|^2 E_n}{\sum_n |c_n|^2} \leq \frac{\sum_n |c_n|^2 E_0}{\sum_n |c_n|^2} = E_0 \quad (9)$$

where Eq. 6 has been utilized. We thus arrive at the central result that makes the variational method possible and practical:

Any approximation to the ground state wavefunction will yield an expectation value of the Hamiltonian that is greater than or equal to the ground state energy. Equality is satisfied only in the case that the approximate wavefunction is also a ground state wavefunction.

In practice, the approximate wavefunction is written in terms of one or more parameters,

$$\phi = \phi(p_1, p_2, \dots, p_N). \quad (10)$$

The set of parameters that yield the best estimate to the ground state energy within the limits of the chosen form of ϕ satisfies

$$\frac{\partial \tilde{E}(p_1, p_2, \dots, p_N)}{\partial p_1} = \frac{\partial \tilde{E}(p_1, p_2, \dots, p_N)}{\partial p_2} = \dots = \frac{\partial \tilde{E}(p_1, p_2, \dots, p_N)}{\partial p_N} = 0 \quad (11)$$

B. Variation with respect to coefficients and Lagrange multipliers

A very important class of electronic structure methods have variational origin. These methods make the initial assumption that the approximate wavefunction is a *sum* of functions satisfying intuitive properties and makes a variation over the expansion coefficients. Let us assume that the approximate wavefunction for a given system may be expanded in terms of a particular set of orbitals. Because we cannot work with an infinitely many number of such orbitals we truncate the sum and just consider the first N terms :

$$\phi(\vec{x}) = \sum_{i=1}^N c_i \chi_i(\vec{x}) \quad (12)$$

For each k , we would like the above expansion to satisfy the minimization condition, i.e.

$$\frac{\partial}{\partial c_k^*} \frac{\langle \phi | \hat{H} | \phi \rangle}{\langle \phi | \phi \rangle}. \quad (13)$$

In addition, we require the approximate wavefunction to remain normalized

$$\langle \phi | \phi \rangle = 1 \quad (14)$$

which then allows us to write rewrite Eq. 13 as

$$\frac{\partial}{\partial c_k^*} \langle \phi | \hat{H} | \phi \rangle = 0 \quad (15)$$

We may in fact, satisfy both Eq. 13 and Eq. 14 by introducing a new quantity

$$K = \langle \phi | \hat{H} | \phi \rangle - \lambda [\langle \phi | \phi \rangle - 1] \quad (16)$$

and extending the minimization property to include the extra parameter λ ,

$$\frac{\partial K}{\partial c_k^*} = \frac{\partial K}{\partial \lambda} = 0 \quad (17)$$

Inserting Eq. 16 into Eq. 17 immediately yields

$$\langle \phi | \phi \rangle - 1 = 0 \quad (18)$$

and proves that minimizing K with respect to *all* of the parameters involved satisfies both of the conditions that we were aiming to satisfy. This method of introducing new variables into the problem to satisfy constraints is often used in classical and quantum mechanics. We may introduce as many variables into the problem as there are constraints. These variables are called *Lagrange multipliers*. The Lagrange multipliers are introduced as arbitrary parameters initially, however we shall see later on that they may correspond to physically meaningful quantities.

Inserting the expansion in Eq. 12 into Eq. 15 yields

$$\begin{aligned} \frac{\partial}{\partial c_k^*} & \left[\langle \sum_i c_i \chi_i | \hat{H} | \sum_j c_j \chi_j \rangle - \lambda \left(\langle \sum_i c_i \chi_i | \sum_j c_j \chi_j \rangle - 1 \right) \right] \\ &= \frac{\partial}{\partial c_k^*} \sum_{i,j} c_i^* c_j [\langle \chi_i | \hat{H} | \chi_j \rangle - \lambda \langle \chi_i | \chi_j \rangle] \\ &= \sum_j c_j [\langle \chi_k | \hat{H} | \chi_j \rangle - \lambda \langle \chi_k | \chi_j \rangle] = 0 \end{aligned} \quad (19)$$

Rearranging Eq. 19, we obtain

$$\sum_j \langle \chi_k | \hat{H} | \chi_j \rangle c_j = \lambda \sum_j \langle \chi_k | \chi_j \rangle c_j \quad (20)$$

which we recognize immediately as a *generalized* eigenvalue equation

$$\mathbf{H} \cdot \mathbf{C} = \lambda \cdot \mathbf{S} \cdot \mathbf{C} \quad (21)$$

where \mathbf{H} and \mathbf{S} are the matrix representations of the Hamiltonian and overlap operators and their elements are defined by.

$$\begin{aligned} H_{nm} &= \langle \chi_n | \hat{H} | \chi_m \rangle \\ S_{nm} &= \langle \chi_n | \chi_m \rangle \end{aligned} \quad (22)$$

λ in Eq. 21 corresponds to a diagonal matrix whose elements on the diagonal are the eigenvalues.

If we use N basis functions to expand the trial function ϕ , Eq. 21 then gives N eigenvalues. But what do the eigenvalues correspond to? In order to see that let's sum both sides of Eq. 20 and isolate λ :

$$\lambda = \frac{\sum_{j,k} c_k^* c_j \langle \chi_k | \hat{H} | \chi_j \rangle}{\sum_{j,k} c_k^* c_j \langle \chi_k | \chi_j \rangle} = \frac{\langle \phi | \hat{H} | \phi \rangle}{\langle \phi | \phi \rangle}. \quad (23)$$

Eq. 23 implies that each of the N eigenvectors correspond to a series of expansion coefficients yielding different ϕ 's and each λ corresponds to a different expectation value. The eigenvector corresponds to the smallest eigenvalue then corresponds to the best ϕ and the smallest eigenvalue itself is the closest approximation to the ground state energy for the approximate form in Eq. 12.

1. An example : Delta-function potential

As discussed above, one can imagine solving a variational problem adopting two different but equivalent strategies :

1. Assume a single wavefunction form having many parameters and minimize over the parameters

2. Assume a linear combination of wavefunctions without any parameters and minimize over the coefficients.

The second strategy is usually preferred in all but the simplest systems, both because the introduction of the variational parameters is more easily controlled and also because it results in an eigenvalue problem that has a straightforward solution. In the example below we are going to compare these two different families of solutions.

Question 1 : A very simple model for the potential seen by a hydrogen electron due to the nucleus is a single delta function $V_H(x) = -\sqrt{\pi}\delta(x)$. Make an estimate to the ground state energy assuming that the approximate wavefunction $\phi_\alpha(x) = \exp(-\alpha x^2/2)$.

Solution : First let us write the one-dimensional Hamiltonian using atomic coordinates ($\hbar = m_e = 1$), which we'll see in detail later.

$$\hat{H} = \underbrace{-\frac{1}{2} \frac{d^2}{dx^2}}_{\hat{T}} + \underbrace{(-\sqrt{\pi}\delta(x))}_{\hat{V}} \quad (24)$$

The calculation of the energy requires the evaluation of a good deal of Gaussian-like integrals. We will therefore first remind ourselves of some simple identities :

$$\int_{-\infty}^{\infty} e^{-\beta x^2} dx = \sqrt{\pi}\beta^{-1/2} \quad (25)$$

$$\int_{-\infty}^{\infty} x^2 e^{-\beta x^2} dx = -\frac{\partial}{\partial \beta} \int_{-\infty}^{\infty} e^{-\beta x^2} dx = -\frac{\partial(\sqrt{\pi}\beta^{-1/2})}{\partial \beta} = \frac{1}{2}\sqrt{\pi}\beta^{-3/2} \quad (26)$$

Another identity that would make our life easier concerns the kinetic energy. For a given wavefunction, the kinetic energy may be written as :

$$T = -\frac{1}{2} \int_{-\infty}^{\infty} \phi dx^*(x) \frac{d}{dx^2} \phi(x) dx = \frac{1}{2} \int_{-\infty}^{\infty} |dx\phi'(x)|^2 \quad (27)$$

This result is obtained by employing integration by parts and taking the resulting surface term to zero.

The potential energy is much more easily calculated since it is just a delta function

$$V = -\sqrt{\pi} \int_{-\infty}^{\infty} \delta dx(x) e^{-\alpha x^2} dx = -\sqrt{\pi} \quad (28)$$

Finally putting everything together, we obtain the α -dependent variational energy as

$$\tilde{E}(\alpha) = \frac{1}{2}\alpha - \alpha^{1/2}. \quad (29)$$

By taking the derivative with respect to α of Eq. 29, we find that the value of α which minimizes this energy is $\alpha = 1$ and the corresponding minimum energy is $E(\alpha = 1) = -\frac{1}{2}$

Question 2 : Consider the same delta potential as in Question 1. This time, make an estimate to the ground state energy assuming that the approximate wavefunction is a linear combination of two Gaussian functions of different width, $\phi(x) = c_1 e^{-x^2/2} + c_2 e^{-2x^2/2}$.

Solution : In accordance with Eq. 21, we need to first construct the the Hamiltonian and overlap matrices for the set of two functions in the linear expansion above, $\phi_1(x) = e^{-x^2/2}$ and $\phi_2(x) = e^{-2x^2/2}$. Let's see how we construct the matrix elements with an example :

$$\begin{aligned} H_{11} &= T_{11} + V_{11} = \langle \chi_1 | \hat{T} | \chi_1 \rangle + \langle \chi_1 | \hat{V} | \chi_1 \rangle \\ &= \frac{1}{2} \int_{-\infty}^{\infty} x^2 e^{-x^2} dx - \int_{-\infty}^{\infty} \sqrt{\pi} \delta(x) e^{-x^2} dx = \frac{1}{2} \frac{\sqrt{\pi}}{2} - \sqrt{\pi} = -\frac{\sqrt{\pi}}{2} \\ S_{11} &= \langle \chi_1 | \chi_1 \rangle = \int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi} \end{aligned} \quad (30)$$

When all the matrix elements are calculated we obtain the following generalized eigenvalue equation

$$\sqrt{\pi} \begin{pmatrix} 1/2 & (2/27)^{1/2} - 1 \\ (2/27)^{1/2} - 1 & \sqrt{1/2} - 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \sqrt{\pi} \begin{pmatrix} 1 & (2/3)^{1/2} \\ (2/3)^{1/2} & \sqrt{1/2} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \quad (31)$$

This generalized eigenvalue equation can be solved using algorithms available in software such as Matlab or may be hand-coded. Below is an excerpt from an Octave (Matlab clone for Linux) session.

```

octave:1> a=(2/27)^(1/2)
a = 0.27217
octave:2> b=(2/3)^(1/2)
b = 0.81650
octave:3> H=[0.5 a-1;
> a-1 sqrt(1/2)-1];
octave:4> H
H =

    0.50000   -0.72783
   -0.72783   -0.29289

octave:5> S=[1 b;
> b sqrt(1/2)];
octave:6> S
S =

    1.00000    0.81650
    0.81650    0.70711

octave:7> [AA,BB,Q,Z,V,W,lambda]=qz(H,S);
octave:8> V
V =

   -0.28422   -0.85324
   -1.00000    1.00000

octave:9> lambda
lambda =

   -0.53212
   31.42433

```

The algorithm used in the above calculation is called *QZ decomposition* and the resulting eigenvalues are contained in the variable `lambda`. The ground state energy is given by the lower of the two eigenvalues, giving an energy slightly smaller than the one found in the first question. By the variational principle discussed in these notes, this energy *has to be* a better approximation to the true ground state energy. The coefficients c_1 and c_2 are contained in the columns of the eigenvector `V` and we see that for the ground state wavefunction, χ_1 gives a much smaller contribution than χ_2 .

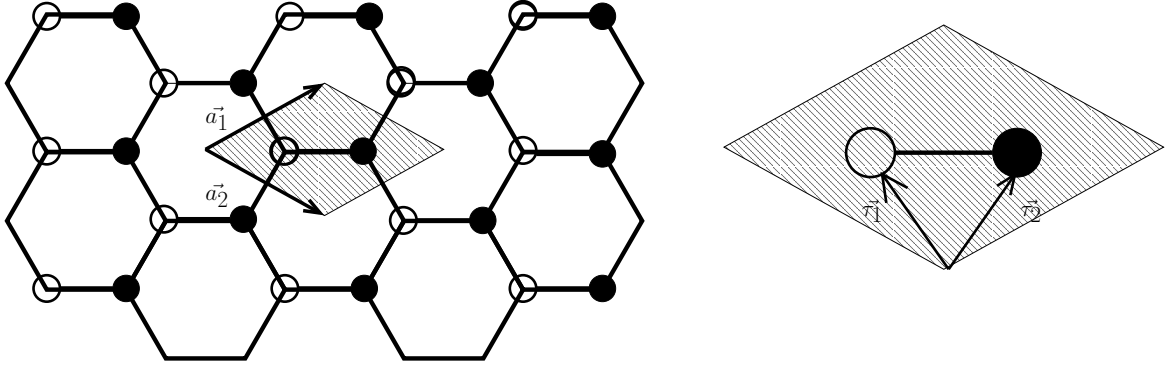
II. CONCEPTS IN PERIODIC SYSTEMS

A. Periodicity and crystals

A *crystal* is a periodic repetition of a set of nuclei in space. It is completely specified by the positions of the nuclei, known as the *basis* in one repeat unit and the translation vectors.

$$\text{Crystal structure} = \text{Bravais lattice} + \text{basis}$$

where the term *Bravais lattice* refers to the lattice generated by the periodic repetition of a single nuclear position.



In the figure above, we see a lattice of black and white atoms. These could be B and N atoms respectively for a BN-sheet. Using one of the equivalent unit cells (hashed area) and combinations of the unit vectors, \vec{a}_1 and \vec{a}_2 , we can generate the entire lattice.

A general translational vector in a periodic lattice can be written as the sum of integer multiples of unit vectors,

$$\vec{T}(\vec{n}) = \vec{T}(n_1, n_2, n_3) = n_1\vec{a}_1 + n_2\vec{a}_2 + n_3\vec{a}_3 \quad (32)$$

where the three unit vectors \vec{a}_i need not be orthonormal. In fact, in the figure above, we see that the two unit vectors spanning the space are not orthonormal.

The lattice vectors do not say anything about how many atoms there are in the unit cell and where they are located. This information is given by the *basis vectors*, which are denoted by \vec{r}_1 and \vec{r}_2 in the above picture. While there are at most three lattice vectors for each unit cells, each unit may contain an arbitrary number of atoms in its basis.

B. Reciprocal space

Solids are obviously finite materials with bounding surfaces. However, for practical purposes, we treat them as infinite systems satisfying specific boundary conditions[1], which impose periodicity in all three dimensions,

$$\Psi(\vec{r}) = \Psi(\vec{r} + N_1\vec{a}_1) = \Psi(\vec{r} + N_2\vec{a}_2) = \Psi(\vec{r} + N_3\vec{a}_3) \quad (33)$$

where N_i are the numbers of unit cells in each of the three directions along the lattice vectors. Such boundary conditions are referred to as *Born-von Karman* boundary conditions.

The long-range periodicity described above is not the only kind of periodicity we have in a crystal. The crystal has another symmetry on a much shorter scale, namely the lattice vectors. Any function defined for a crystal, such as the electron density, is bound to be periodic, repeating itself with the same translation vectors as those that span the lattice. Thus,

$$f(\vec{r} + \vec{T}(n_1, n_2, n_3)) = f(\vec{r}). \quad (34)$$

where $\vec{T} = n_1\vec{a}_1 + n_2\vec{a}_2 + n_3\vec{a}_3$ is a translation vector.

Such periodic functions lend themselves easily to Fourier transform. Under certain conditions, it becomes more advantageous to deal with the Fourier components of such systems rather dealing with them in real space. The forward Fourier transform is conventionally defined as

$$f(\vec{q}) = \frac{1}{\Omega_{\text{crystal}}} \int_{\Omega_{\text{crystal}}} d\vec{r} f(\vec{r}) \exp(i\vec{q} \cdot \vec{r}). \quad (35)$$

where Ω_{crystal} is the volume of the crystal. The Born-von Karman conditions set constraints on the allowed wavevectors, \vec{q} , which we may use while determining the Fourier components. This may easily be seen by adding a translation vector, $N_i\vec{a}_i$ to \vec{r} in Eq. 33 and making use of Eq. 35,

$$f(\vec{q}) = \frac{1}{\Omega_{\text{crystal}}} \int_{\Omega_{\text{crystal}}} d\vec{r} f(\vec{r} + N_1\vec{a}_1) \exp(i\vec{q} \cdot (\vec{r} + N_1\vec{a}_1)) = \frac{\exp(i\vec{q} \cdot (N_1\vec{a}_1))}{\Omega_{\text{crystal}}} \int_{\Omega_{\text{crystal}}} d\vec{r} f(\vec{r}) \exp(i\vec{q} \cdot \vec{r}) \quad (36)$$

Comparing Eq. 35 and Eq. 36, we see that $\exp(i\vec{q} \cdot (N_1\vec{a}_1)) = 1$. The same argument may be applied to the directions of the other two lattice vectors, yielding the following restriction on the wavevectors \vec{q} .

$$\vec{q} \cdot \vec{a}_i = \frac{2\pi n_i}{N_i} \quad \text{where } n_i = 0, 1, 2, \dots, N_i - 1 \quad (37)$$

We thus have the familiar situation that confinement causes the wavevectors to be quantized (remember particle in a box). Next, we make use of the periodicity on the scale of the lattice constants. For a function f having this sort of periodicity, Eq. 35 may be written as

$$\begin{aligned}
 f(\vec{q}) &= \frac{1}{\Omega_{\text{crystal}}} \sum_{n_1, n_2, n_3} \int_{\Omega_{\text{cell}}} d\vec{r} f(\vec{r}) \exp(i\vec{q} \cdot \vec{r}) \\
 &= \frac{1}{N_{\text{cell}}} \sum_{n_1, n_2, n_3} \exp[\vec{q} \cdot \vec{T}(n_1, n_2, n_3)] \frac{1}{\Omega_{\text{cell}}} \int d\vec{r} f(\vec{r}) \exp(i\vec{q} \cdot \vec{r}) \\
 &= \frac{1}{N_{\text{cell}}} \prod_i \sum_{n_i} \exp[\vec{q} \cdot (n_i \vec{a}_i)] \frac{1}{\Omega_{\text{cell}}} \int d\vec{r} f(\vec{r}) \exp(i\vec{q} \cdot \vec{r})
 \end{aligned} \tag{38}$$

where we make use of the fact that the integral over the volume of the crystal can be broken down into integrals of identical crystals. Consider one of the sums over $\{n_i\}$:

$$\sum_{n_i=0}^{N_i-1} (e^{i\vec{q} \cdot \vec{a}_i})^{n_i} = \frac{1 - e^{2\pi i \frac{\text{integer}}{N_i} N_i}}{1 - e^{2\pi i \frac{\text{integer}}{N_i}}} \tag{39}$$

by Eq. 33 and the use of a geometric series. Eq. 39 gives zero unless

$$e^{\vec{q} \cdot \vec{a}_i} = 1 \Rightarrow \vec{q} \cdot \vec{a}_i = 2\pi m \tag{40}$$

where m is an integer. Thus, periodicity at this small scale implies a new selection rule for the allowed wavevectors, \vec{q} . The space spanned by the allowed, discrete set of $\{\vec{q}\}$ that satisfy Eq. 40 is called the *reciprocal lattice*. They are referred to as a lattice because just like lattice vectors, they define a regular array of atoms with a well-defined periodicity. Any point in the reciprocal lattice can be expressed in terms of a set of minimal set of \vec{q} 's which are defined by

$$\vec{b}_i \cdot \vec{a}_j = 2\pi \delta_{ij}. \tag{41}$$

An arbitrary point on the lattice is then given by

$$\vec{G}(m_1, m_2, m_3) = m_1 \vec{b}_1 + m_2 \vec{b}_2 + m_3 \vec{b}_3. \tag{42}$$

One way to satisfy Eq. 41 is to construct \vec{b}_i 's such that they obey

$$\vec{b}_i = 2\pi \frac{\vec{a}_j \times \vec{a}_k}{\vec{a}_i \cdot (\vec{a}_j \times \vec{a}_k)} \tag{43}$$

The term on the denominator is numerically equal to the volume of the unit cell in real space. For each \vec{G} in Eq. 42, the Fourier transform of the periodic function can then be written as

$$f(\vec{G}) = \frac{1}{\Omega_{\text{cell}}} \int_{\Omega_{\text{cell}}} d\vec{r} f(\vec{r}) e^{i\vec{G} \cdot \vec{r}}. \tag{44}$$

1. An example : Face-centered cubic lattice and its Brillouin zone

A *face-centered cubic* (fcc) lattice is one where the atoms are located on the corners of a cube as well as on the face diagonals. The primitive unit cell is thus given by the following lattice vectors :

$$\begin{aligned}
 \vec{a}_1 &= a\left(\frac{1}{2}, \frac{1}{2}, 0\right) \\
 \vec{a}_2 &= a\left(\frac{1}{2}, 0, \frac{1}{2}\right) \\
 \vec{a}_3 &= a\left(0, \frac{1}{2}, \frac{1}{2}\right)
 \end{aligned} \tag{45}$$

Using the definition in Eq. 43, we find that the reciprocal lattice vectors are

$$\begin{aligned}\vec{b}_1 &= \frac{2\pi}{a}(1, 1, -1) \\ \vec{b}_2 &= \frac{2\pi}{a}(1, -1, 1) \\ \vec{b}_3 &= \frac{2\pi}{a}(-1, 1, 1)\end{aligned}\quad (46)$$

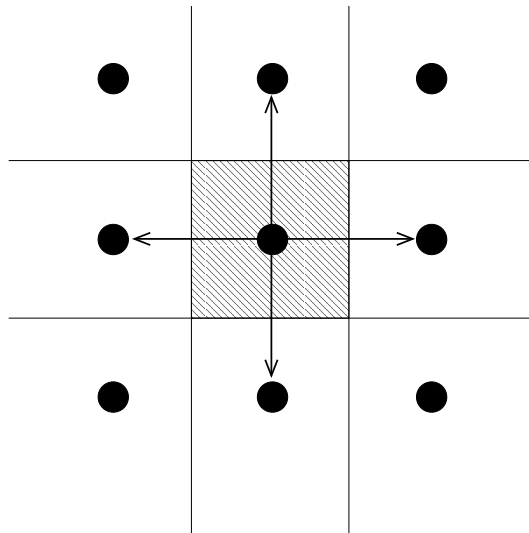
Note that the reciprocal vectors that belong to the fcc lattice define a *body-centered cubic* (bcc) lattice in the reciprocal space.

C. The Brillouin Zone

From Eq. 44, one easily sees that there is a certain periodicity also in the reciprocal space, that is

$$f(\vec{G} + m_i \vec{b}_i) = f(\vec{G}). \quad (47)$$

Thus it makes sense to define a *unit cell* also in reciprocal space beyond which f repeats itself. Such a unit cell is already widely used in literature and it is called the *Brillouin zone*. The *Brillouin zone* is defined by the area surrounded by the planes that are perpendicular bisectors of the vectors from the origin to the reciprocal lattice points.



D. Bloch's theorem

In a single-electron picture, the Hamiltonian is invariant under lattice translations. Thus the Hamiltonian commutes with the translation operator

$$[\hat{H}, \hat{T}_{\vec{n}}] = 0 \quad \text{where} \quad \hat{T} = \vec{n}_1 \vec{a}_1 + \vec{n}_2 \vec{a}_2 + \vec{n}_3 \vec{a}_3. \quad (48)$$

From elementary quantum mechanics, remember that operators that commute share a common set of eigenfunctions. Such a common set may thus be found also for \hat{H} and $\hat{T}_{\vec{n}}$. Let's consider in particular, $\vec{n}_1 = (1, 0, 0)$. The Born-von Karman conditions require that the eigenfunctions remain unchanged if the translation operator is applied N_1 times

$$\hat{T}_{\vec{n}_1} \psi(\vec{r}) = (t_{\vec{n}_1})^{N_1} \psi(\vec{r}) = \psi(\vec{r}) \quad (49)$$

where $t_{\vec{n}_1}$ is the eigenvalue of the translation operator $\hat{T}_{\vec{n}_1} \psi(\vec{r})$. The requirement

$$(t_{\vec{n}_1})^{N_1} = 1 \quad (50)$$

is satisfied for e^{ik} where k is, at this point, just a real number. Eq. 50, on the other hand, dictates that

$$e^{ikN_1} = 1 \quad \Rightarrow \quad kN_1 = 2\pi n_1 \quad \Rightarrow \quad k = \frac{2\pi n_1}{N_1} \quad (51)$$

The same relation may be written by a general k restricted to the first Brillouin zone

$$\vec{k} = \frac{n_1}{N_1}\vec{b}_1 + \frac{n_2}{N_2}\vec{b}_2 + \frac{n_3}{N_3}\vec{b}_3 \quad (52)$$

We finally arrive at Bloch's theorem which states that the eigenfunctions of a lattice-periodic Hamiltonian must satisfy

$$\hat{T}_{\vec{n}}\psi(\vec{r}) = e^{i\vec{k}\cdot\vec{T}_{\vec{n}}}\psi(\vec{r}). \quad (53)$$

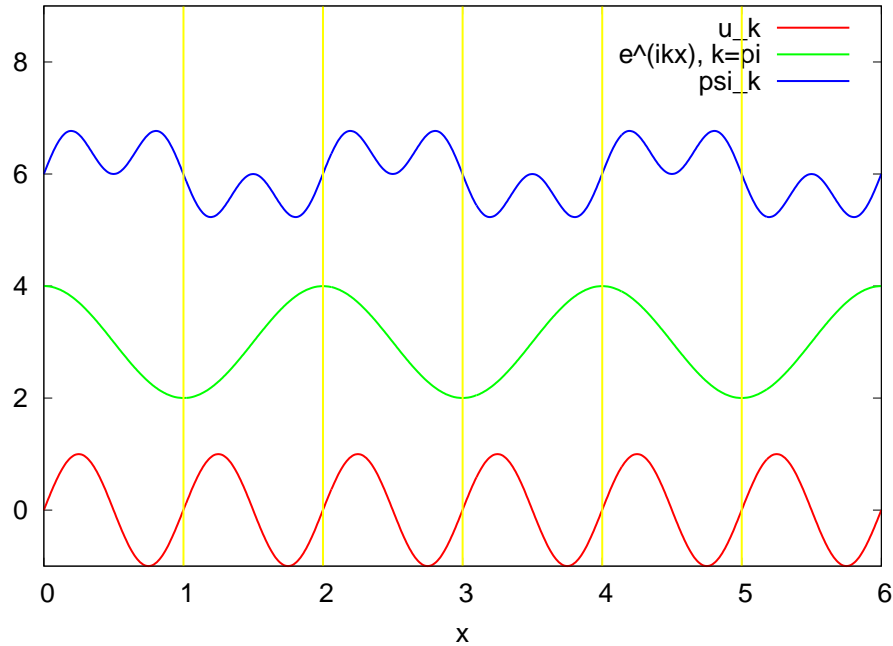
Bloch's theorem states that eigenstates of a periodic Hamiltonian pick up a phase when translated by a translation vector and may thus be labeled by the wavevector that characterizes the phase factor. In particular, Eq. 53 by wavefunctions written as

$$\psi_{\vec{k}}(\vec{r}) = e^{i\vec{k}\cdot\vec{r}}u_{\vec{k}}(\vec{r}) \quad (54)$$

where $u_{\vec{k}}(\vec{r})$ is a lattice periodic function with

$$u_{\vec{k}}(\vec{r} + \vec{T}_{\vec{n}}) = u_{\vec{k}}(\vec{r}). \quad (55)$$

Eq. 54 is an equivalent statement of Bloch's theorem. Let's illustrate this with the following picture in one dimension.



The bottommost line is a lattice periodic wavefunction of period $a = 1$ in some units. The line in the middle is the phase factor e^{ikx} for $k = \pi/2a$ and the topmost curve is the wavefunction that is their product (its real part to be more precise). As you see, unlike the electron density, the wavefunction *need not* have the periodicity of the lattice. It is still required however to obey the Born-von Karman boundary conditions.

The ability provided by the Bloch theorem to break down the wavefunction into a lattice-periodic function $u_{\vec{k}}$ and a phase factor allows us to work with a reduced, wavevector-dependent Hamiltonian. Applying the real Hamiltonian \hat{H} of the system on to the wavefunction given in Eq. 54, we have[2]

$$\begin{aligned} \hat{H}\psi_{i,\vec{k}}(\vec{r}) &= \epsilon_{i,\vec{k}}\psi_{i,\vec{k}}(\vec{r}) \\ \hat{H}e^{i\vec{k}\cdot\vec{r}}u_{i,\vec{k}}(\vec{r}) &= \epsilon_{i,\vec{k}}e^{i\vec{k}\cdot\vec{r}}u_{i,\vec{k}}(\vec{r}) \end{aligned} \quad (56)$$

Multiplying the second line in Eq. 56 by $e^{-i\vec{k}\cdot\vec{r}}$, we have

$$\begin{aligned} e^{-i\vec{k}\cdot\vec{r}}\hat{H}e^{i\vec{k}\cdot\vec{r}}u_{i,\vec{k}}(\vec{r}) &= \epsilon_{i,\vec{k}}u_{i,\vec{k}}(\vec{r}) \\ \hat{H}_{\vec{k}}u_{i,\vec{k}}(\vec{r}) &= \epsilon_{i,\vec{k}}u_{i,\vec{k}}(\vec{r}) \end{aligned} \quad (57)$$

where $\hat{H}_{\vec{k}} = e^{-i\vec{k}\cdot\vec{r}}\hat{H}e^{i\vec{k}\cdot\vec{r}}$. It is therefore possible to scan the Brillouin zone by writing and solving a new Hamiltonian for each wavevector. Each $\hat{H}_{\vec{k}}$ will yield a spectrum of eigenvalues and in the limit of large crystal dimension, these eigenvalues will merge into a continuous graphical representation of eigenvalues, which is called the *band structure*. The band structure is a very important tool for deciding upon the electronic properties of a crystal.

1. An example : Tight-binding model for atoms on a chain

In order to illustrate the arguments we have introduced above, let's see a very simple one-dimensional example [3]. Assume that we have a one-dimensional arrangement of infinitely many atoms on a chain and assume Born-van Karman boundary conditions with a period of N atoms.

We are going to employ what is known as the *tight-binding model* where each wavefunction of the whole crystal can be expanded as a linear combination of atomic orbitals located on each of the atoms.

$$\psi(\vec{r}) = \sum_{n,l} c_{n,l}\phi_l(\vec{r} - \vec{R}_n) \quad (58)$$

where $\{\phi_l(\vec{r})\}$ could refer to a single orbital or a collection of orbitals, denoted by angular momentum orbital number l . When dealing with silicon, for instance, it is sufficient to include only the s ($l=0$) and p ($l=1$) orbitals, since they are the only ones that participate in bonding, while describing a Cu crystal would require the inclusion of d ($l=2$) orbitals as well. For our simple example, we will only consider the s -orbitals as well. We then switch to the matrix representation of \hat{H} and hS in the atomic orbital basis

$$\begin{aligned} H_{nm} &= \langle\phi_n|\hat{H}|\phi_m\rangle \\ S_{nm} &= \langle\phi_n|\phi_m\rangle \end{aligned} \quad (59)$$

where now the index n represents the atom where the orbital is located and the orbital angular momentum has been dropped. Furthermore, we also assume that the orbitals that are located on different atoms do not overlap sufficiently so that the overlap operator hS takes the form of identity in the representation in Eq. 59.

In tight-binding calculations, the matrix elements above is not actually calculated but is *parametrized*. The parameters are fitted to reproduce such material properties as lattice constants and elastic moduli. We will assume the following over-simplified, *second nearest neighbor model* :

$$H_{nm} = \begin{cases} t & \text{if } n = m \\ h_1 & \text{if } n = m \mp 1 \\ h_2 & \text{if } n = m \mp 2 \\ 0 & \text{otherwise} \end{cases} \quad (60)$$

where $h_1 > h_2$. This means that the interaction is only confined to the neighbors that are separated by at most two lattice sites. The final point to remember is that since we assume a periodicity of N , the atoms labeled 1 and N are regarded as immediate neighbors and atoms labeled 2 and N are regarded as second nearest neighbors (same argument goes for pairs 1- $N-1$ and 2- N).

As an example, for $t = -1.0$, $h_1 = -0.5$ and $h_2 = -0.3$, the Hamiltonian matrix is given by

$$H = \begin{pmatrix} -1.0 & -0.5 & -0.3 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & -0.3 & -0.5 \\ -0.5 & -1.0 & -0.5 & -0.3 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & -0.3 \\ -0.3 & -0.5 & -1.0 & -0.5 & -0.3 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & -0.3 & -0.5 & -1.0 & -0.5 & -0.3 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & -0.3 & -0.5 & -1.0 & -0.5 & -0.3 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & -0.3 & -0.5 & -1.0 & -0.5 & -0.3 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 & -0.3 & -0.5 & -1.0 & -0.5 & -0.3 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & -0.3 & -0.5 & -1.0 & -0.5 & -0.3 \\ -0.3 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & -0.3 & -0.5 & -1.0 & -0.5 \\ -0.5 & -0.3 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & -0.3 & -0.5 & -1.0 \end{pmatrix} \quad (61)$$

where the periodic pattern is easily recognized. The eigenvalues of this matrix are

$$\Lambda = (-2.60 \quad -1.99 \quad -1.99 \quad -0.82 \quad -0.82 \quad -0.60 \quad -0.38 \quad -0.38 \quad -0.21 \quad -0.21) \quad (62)$$

and give the electronic spectra of the one-dimensional chain. The corresponding eigenvectors give the coefficients in Eq. 58.

For a system of 10 atoms, a 10×10 matrix can easily be dealt with. But for larger systems and more complicated interactions, it is imperative to utilize the periodicity of the system via Bloch's theorem. To this end, let's rewrite the wavefunction in Eq. 58 in the Bloch form

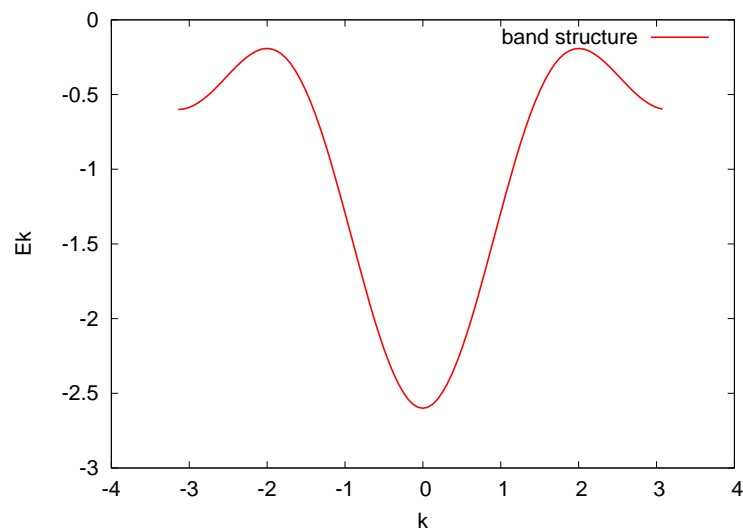
$$\psi_k(\vec{r}) = \sum_n e^{ikn} \phi(\vec{r} - \vec{R}_n) \quad (63)$$

where the $k = 0, \frac{2\pi}{N}, \frac{4\pi}{N} \dots \frac{(N-2)\pi}{N}$. It is then easy to prove (see Springborg) that the Hamiltonian is diagonalized by this form of the wavefunction and moreover, the overlap matrix is still unity. This implies that the eigenvalue for each k is simply given by

$$\varepsilon_k = \sum_n e^{ik(n-1)} \langle \phi_n | H | \phi_1 \rangle \quad (64)$$

which yields the identical spectrum as displayed above in Eq. 62.

Obtaining the eigenvalues as a function of k gives us the opportunity to plot the eigenvalues as a function of k . This is precisely the *band structure* that was mentioned earlier.



The above example illustrates the recipe for going from the full-dimensional problem, which is often intractably large to a problem that has a much more manageable size, namely a single unit cell. In the present case, there's only a single atom in the unit cell and therefore we obtain a single band in our band structure. In more complicated problems with several atoms in the unit cell, the band structure has as many bands as there are atoms and is therefore much more complicated[4].

[1] This is a good approximation unless the system size is so small that surface effects cannot be neglected

[2] The new index i designates each eigenfunction in the spectrum of a given \vec{k}

[3] Adapted from Springborg

[4] For the interested reader, the Octave code relevant to this problem can be found on the course Web page