

## PHYS 741 : Principles of Density Functional Theory

- **Credits :** 3
- **Course Web Site :** <http://www.physics.metu.edu.tr/~hande/teaching.html>
- **Grading :**
  - Homework (50%) : 5 or 6 homework assignments
  - Oral exam (20%) : Mid-semester
  - Written exam (20%) : At the end of the semester
  - Term paper (10%) : At the end of the semester (subjects are announced on the course Web site)
- **Term paper :** The term paper will be a literature survey of at least 10 pages, the subject of which is to be chosen by the students from the list provided on the Web site. The paper MUST be written in the students' own words, i.e. NO COPY AND PASTE; otherwise points will be taken away. The paper must be a coherent treatment of the subject involving not one but several references, properly cited. Suggestions will be provided by the instructor upon the students' request. The students may choose their subjects three weeks-a month into the semester.
- **Lectures :** 3-hour lecture Thursdays 13:40-16:30 in the Seminar room.
- **Textbooks:**
  - *Methods of Electronic Structure Calculations*, Micheal Springborg
  - *Electronic Structure : Basic Theory and Practical Methods*, Richard Martin
  - *Density-functional theory of atoms and molecules*, Robert Parr and Weitao Yang
- **Other references :**
  - *A Bird's Eye View of Density-Functional Theory*, Klaus Capelle, arxiv:cond-mat/0211443
  - *Iterative minimization techniques for ab initio total-energy calculations : molecular dynamics and conjugate gradients*, Michael C. Payne *et al.*, Rev Mod Phys, **64**, 1045 (1992)
  - *The ABC of DFT*, Kieron Burke , <http://dft.rutgers.edu/kieron/beta>
  - *Pseudopotential methods in condensed matter applications*, Warren Pickett, Comp Phys Rep, **9**, 115-197 (1989)
  - Various papers to be assigned during class
- **Topics to be Covered :**
  1. Preliminary topics : variational principle, solid state concepts, the many-body Hamiltonian, Born-Oppenheimer approximation
  2. The Hartree-Fock approach
  3. Hamiltonian in terms of density, Hohenberg-Kohn theorems, Kohn-Sham formalism
  4. Exchange and correlations : LDA and GGA approximations
  5. Pseudopotentials

6. Plane-wave formalism
  7. Tricks for self-consistent solution of the Kohn-Sham system : mixing and diagonalization techniques
  8. Forces, stress and the Hellman-Feynman theorem
  9. How to deal with the ions : Ewald sum
  10. Practical matters and application of DFT with examples, PWSCF
  11. Specialized topics (time permitting)
    - Car-Parrinello molecular dynamics
    - Exchange-correlation functionals
    - Density functional perturbation theory
    - Beyond DFT : TDDFT, GW, chemical methods
- **Practice** : There will be two computer-related exercises in this course :
    1. A DFT code that we will develop from scratch using Octave or Matlab. We will hold a small introductory tutorial on Octave in the beginning of the course for those who are not familiar with this. This will be done after-hours.
    2. An open-source DFT code that we will use frequently — Quantum-Espresso. We will designate a few of our lecture hours as exercise in this code to highlight some ideas and we will also use it for homework assignments.
  - **Things you will need** : You will need some hardware/software items while doing the homework and/or participating in the class :
    1. A computer (preferably laptop) which has Linux (I suggest Ubuntu, but any distro is OK.). In order to participate in the tutorials etc, you will occasionally need to bring in your laptop.
    2. Quantum-Espresso v. 5.0 (<http://www.quantum-espresso.org/>), compiled on your computer in the serial mode.
    3. XCrySDen (X-window crystalline structures and densities) visualization program v. 1.5.53 (<http://www.xcrysden.org/>). You need the **semishared** version. This is just a download — no installation necessary.
    4. Sometimes linux distributions do not come with Octave. If you are using Ubuntu, after you have installed it, install also the latest version of Octave.

# Lecture I : The variational principle

## I. 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  $\psi_0$  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  $\psi_0$

Unless we are very lucky, the approximate wavefunction  $\phi$  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  $\psi$  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,  $\phi$ . 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  $\phi$  satisfies

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

## II. VARIATION WITH RESPECT TO COEFFICIENTS (LINEAR VARIATION) 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 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  $\lambda$ ,

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

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.

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

$\lambda$  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  $\phi$ , 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  $\lambda$  :

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

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

### III. 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 :

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

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

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

Finally putting everything together, we obtain the  $\alpha$ -dependent variational energy as

$$\tilde{E}(\alpha) = \frac{1}{4}\alpha - \alpha^{1/2}. \quad (29)$$

By taking the derivative with respect to  $\alpha$  of Eq. 29, we find that the value of  $\alpha$  which minimizes this energy is  $\alpha = 4$  and the corresponding minimum energy is  $E(\alpha = 1) = -1$ .  $\tilde{E}(\alpha)$  as a function of  $\alpha$  is displayed in Fig. 2

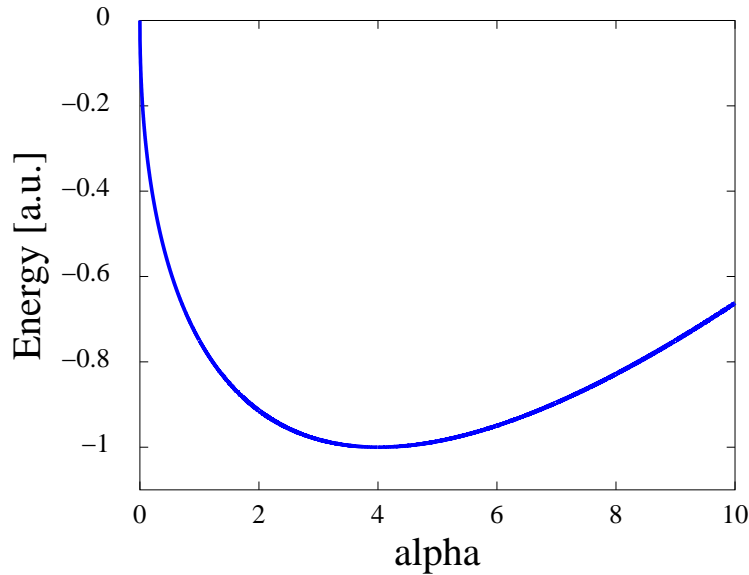


FIG. 1:  $\tilde{E}$  as a function of  $\alpha$ . The energy minimum is clearly seen at  $\alpha = 4$ .

**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 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 :

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

When all the matrix elements are calculated we obtain the Hamiltonian and overlap matrices in the basis of  $\{\chi_1, \chi_2\}$

$$H = \sqrt{\pi} \begin{pmatrix} -3/4 & (2/27)^{1/2} - 1 \\ (2/27)^{1/2} - 1 & \sqrt{1/8} - 1 \end{pmatrix} \quad (31)$$

$$S = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \sqrt{\pi} \begin{pmatrix} 1 & (2/3)^{1/2} \\ (2/3)^{1/2} & \sqrt{1/2} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \quad (32)$$

Using these matrices, the generalized eigenvalue equation in Eq. 21 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=sqrt(2/27)
a = 0.27217
octave:2> b=sqrt(2/3)
b = 0.81650
octave:3> H=sqrt(pi)*[-3/4 a-1;a-1 sqrt(1/8)-1]
H =

    -1.3293    -1.2901
    -1.2901    -1.1458

octave:4> S=sqrt(pi)*[1 b;b sqrt(1/2)]
S =

    1.7725    1.4472
    1.4472    1.2533

octave:5> [AA,BB,Q,Z,V,W,lambda]=qz(H,S);
octave:6> V
V =

    0.13024   -0.87543
   -1.00000    1.00000

octave:7> lambda
lambda =

   -0.91825
    1.20935
```

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. The coefficients  $c_1$  and  $c_2$  are the elements of the first column of `V`. The other column corresponds to the "excited state" of the variational problem. One must exercise caution in talking about the excited state here because although the ground state energy found using variational principle is an upper bound to the true ground state energy, there is no such guarantee for higher energy states.

The "ground" and "excited" states of this problem are plotted in Fig. ???. The Octave codes used to generate this plot can be found on the course Web site. Note that the integrals and derivatives in the Octave functions have been evaluated numerically rather than analytically.

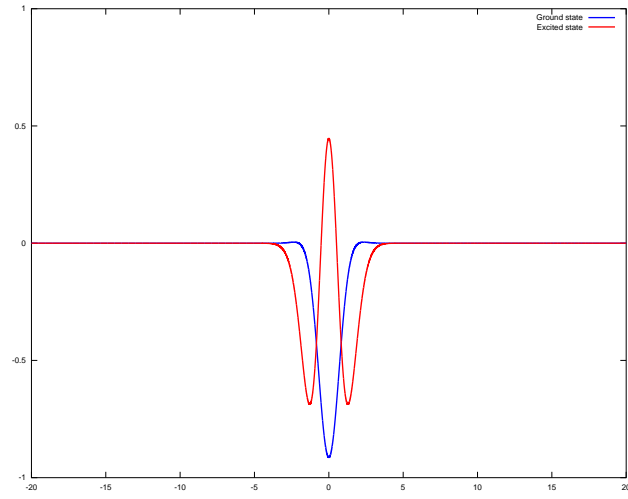


FIG. 2: Ground and excited states of the linear variation problem.



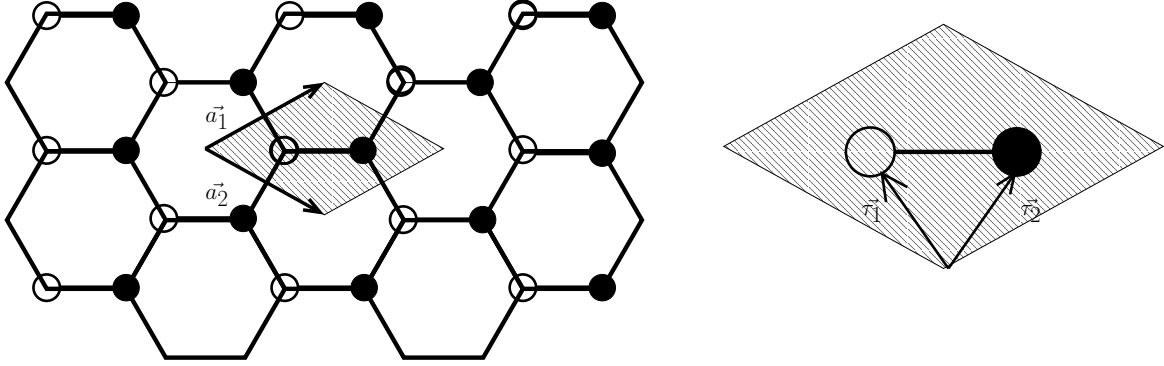
## Lecture II : Fundamental concepts in solid state

### I. 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.

Crystal structure = Bravais lattice + 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,

$$\vec{T}(\vec{n}) = \vec{T}(n_1, n_2, n_3) = n_1\vec{a}_1 + n_2\vec{a}_2 + n_3\vec{a}_3 \quad (1)$$

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 cell, each unit may contain an arbitrary number of atoms in its basis.

### II. 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,

$$\Psi(\vec{r}) = \Psi(\vec{r} + N_1\vec{a}_1) = \Psi(\vec{r} + N_2\vec{a}_2) = \Psi(\vec{r} + N_3\vec{a}_3) \quad (2)$$

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,

$$f(\vec{r} + \vec{T}(n_1, n_2, n_3)) = f(\vec{r}). \quad (3)$$

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

$$f(\vec{q}) = \frac{1}{\Omega_{\text{crystal}}} \int_{\Omega_{\text{crystal}}} d\vec{r} f(\vec{r}) \exp(i\vec{q} \cdot \vec{r}). \quad (4)$$

where  $\Omega_{\text{crystal}}$  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. 2 and making use of Eq. 4,

$$f(\vec{q}) = \frac{1}{\Omega_{\text{crystal}}} \int_{\Omega_{\text{crystal}}} d\vec{r} f(\vec{r} + N_1 \vec{a}_1) \exp(i\vec{q} \cdot (\vec{r} + N_1 \vec{a}_1)) = \frac{\exp(i\vec{q} \cdot (N_1 \vec{a}_1))}{\Omega_{\text{crystal}}} \int_{\Omega_{\text{crystal}}} d\vec{r} f(\vec{r}) \exp(i\vec{q} \cdot \vec{r}) \quad (5)$$

Comparing Eq. 4 and Eq. 5, 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  $\vec{q}$ .

$$\vec{q} \cdot \vec{a}_i = \frac{2\pi n_i}{N_i} \quad \text{where} \quad n_i = 0, 1, 2, \dots, N_i - 1 \quad (6)$$

We thus have the familiar situation that periodicity causes the wavevectors to be quantized. Next, we make use of the periodicity on the scale of the lattice constants. For a function  $f$  having this sort of periodicity, Eq. 4 may be written as

$$\begin{aligned} f(\vec{q}) &= \frac{1}{\Omega_{\text{crystal}}} \sum_{n_1, n_2, n_3} \int_{\Omega_{\text{cell}}} d\vec{r} f(\vec{r} + \vec{T}(n_1, n_2, n_3)) \exp[i\vec{q} \cdot (\vec{r} + \vec{T}(n_1, n_2, n_3))] \\ &= \frac{1}{N_{\text{cell}}} \sum_{n_1, n_2, n_3} \exp[i\vec{q} \cdot \vec{T}(n_1, n_2, n_3)] \frac{1}{\Omega_{\text{cell}}} \int d\vec{r} f(\vec{r}) \exp(i\vec{q} \cdot \vec{r}) \\ &= \frac{1}{N_{\text{cell}}} \prod_i \sum_{n_i} \exp[i\vec{q} \cdot (n_i \vec{a}_i)] \frac{1}{\Omega_{\text{cell}}} \int d\vec{r} f(\vec{r}) \exp(i\vec{q} \cdot \vec{r}) \end{aligned} \quad (7)$$

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  $\{n_i\}$ :

$$\sum_{n_i=0}^{N_i-1} (e^{i\vec{q} \cdot \vec{a}_i})^{n_i} = \frac{1 - e^{2\pi i \frac{\text{integer}}{N_i} N_i}}{1 - e^{2\pi i \frac{\text{integer}}{N_i}}} \quad (8)$$

by Eq. 2 and the use of a geometric series. Eq. 8 gives zero unless

$$e^{i\vec{q} \cdot \vec{a}_i} = 1 \Rightarrow \vec{q} \cdot \vec{a}_i = 2\pi m \quad (9)$$

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. 9 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

$$\vec{b}_i \cdot \vec{a}_j = 2\pi \delta_{ij}. \quad (10)$$

An arbitrary point on the lattice is then given by

$$\vec{G}(m_1, m_2, m_3) = m_1 \vec{b}_1 + m_2 \vec{b}_2 + m_3 \vec{b}_3 = \sum_j m_j \vec{b}_j. \quad (11)$$

which satisfies

$$\vec{G} \cdot \vec{a}_i = 2\pi m_i \quad (12)$$

One way to satisfy Eq. 10 is to construct  $\vec{b}_i$ 's such that they obey

$$\vec{b}_i = 2\pi \frac{\vec{a}_j \times \vec{a}_k}{\vec{a}_i \cdot (\vec{a}_j \times \vec{a}_k)} \quad (13)$$

The term on the denominator is numerically equal to the volume of the unit cell in real space. Rewriting Eq. 4 for valid  $\{\vec{G}\}$  as defined in Eq. 11 yields

$$\begin{aligned} f(\vec{G}) &= \frac{1}{N_{cell}} \prod_i \sum_{n_i} \frac{1}{\Omega_{cell}} \int_{\Omega_{cell}} d\vec{r} f(\vec{r}) e^{i\vec{G} \cdot \vec{r}} \\ &= \frac{1}{\Omega_{cell}} \int_{\Omega_{cell}} d\vec{r} f(\vec{r}) e^{i\vec{G} \cdot \vec{r}}. \end{aligned} \quad (14)$$

#### A. 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 :

$$\begin{aligned} \vec{a}_1 &= a\left(\frac{1}{2}, \frac{1}{2}, 0\right) \\ \vec{a}_2 &= a\left(\frac{1}{2}, 0, \frac{1}{2}\right) \\ \vec{a}_3 &= a\left(0, \frac{1}{2}, \frac{1}{2}\right) \end{aligned} \quad (15)$$

Using the definition in Eq. 13, we find that the reciprocal lattice vectors are

$$\begin{aligned} \vec{b}_1 &= \frac{2\pi}{a}(1, 1, -1) \\ \vec{b}_2 &= \frac{2\pi}{a}(1, -1, 1) \\ \vec{b}_3 &= \frac{2\pi}{a}(-1, 1, 1) \end{aligned} \quad (16)$$

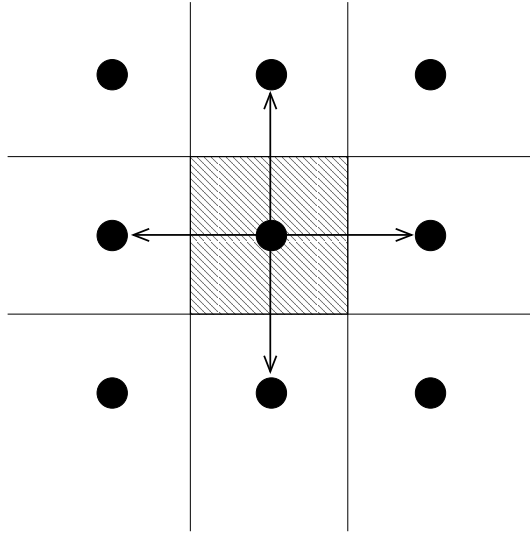
Note that the reciprocal vectors that belong to the fcc lattice define a *body-centered cubic* (bcc) lattice in the reciprocal space.

#### B. The Brillouin Zone

From Eq. 14, one easily sees that there is a certain periodicity also in the reciprocal space, that is

$$f(\vec{G} + m_i \vec{b}_i) = f(\vec{G}). \quad (17)$$

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.



### C. Bloch's theorem

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

$$[\hat{H}, \hat{T}_{\vec{n}}] = 0 \quad \text{where} \quad \hat{T} = \vec{n}_1 \vec{a}_1 + \vec{n}_2 \vec{a}_2 + \vec{n}_3 \vec{a}_3. \quad (18)$$

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,  $\vec{n} = (1, 0, 0)$ . The Born-von Karman conditions require that the eigenfunctions remain unchanged if the translation operator is applied  $N_1$  times

$$\hat{T}_{\vec{n}} \psi(\vec{r}) = (t_{\vec{n}})^{N_1} \psi(\vec{r}) = \psi(\vec{r}) \quad (19)$$

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

$$(t_{\vec{n}})^{N_1} = 1 \quad (20)$$

is satisfied for  $e^{ik}$  where  $k$  is, at this point, just a real number. Eq. 20, on the other hand, dictates that

$$e^{ikN_1} = 1 \quad \Rightarrow \quad kN_1 = 2\pi n_1 \quad \Rightarrow \quad k = \frac{2\pi n_1}{N_1} \quad (21)$$

The same relation may be written by a general, three dimensional  $\vec{k}$  restricted to the first Brillouin zone

$$\vec{k} = \frac{n_1}{N_1} \vec{b}_1 + \frac{n_2}{N_2} \vec{b}_2 + \frac{n_3}{N_3} \vec{b}_3 \quad (22)$$

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

$$\hat{T}_{\vec{n}} \psi(\vec{r}) = e^{i\vec{k} \cdot \vec{T}_{\vec{n}}} \psi(\vec{r}). \quad (23)$$

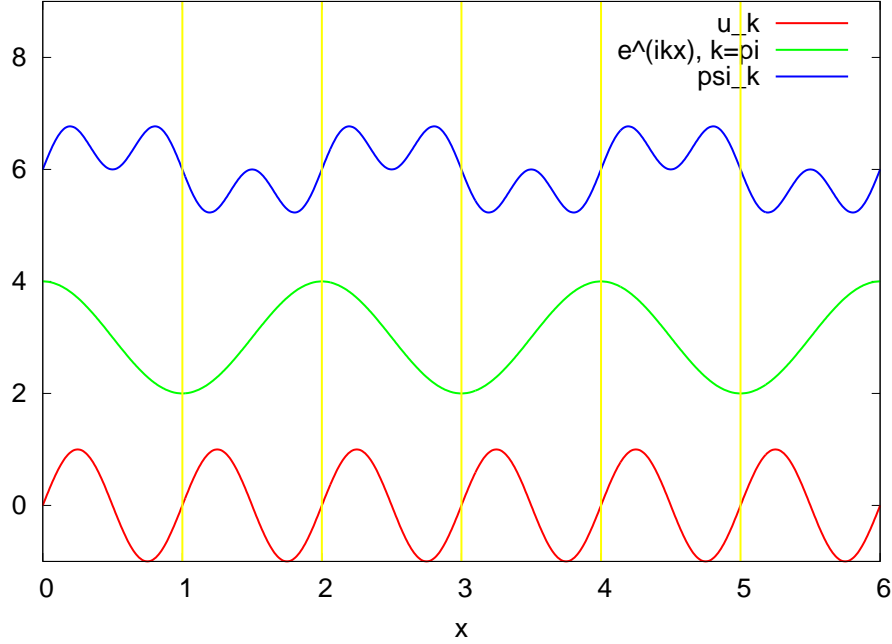
Bloch's theorem states that eigenstates of a periodic Hamiltonian pick up a phase when translated by a translation vector and may thus be labeled by the wavevector that characterizes the phase factor. In particular, Eq. 23 by wavefunctions written as

$$\psi_{\vec{k}}(\vec{r}) = e^{i\vec{k} \cdot \vec{r}} u_{\vec{k}}(\vec{r}) \quad (24)$$

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

$$u_{\vec{k}}(\vec{r} + \vec{T}_{\vec{n}}) = u_{\vec{k}}(\vec{r}). \quad (25)$$

Eq. 24 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. 31, we have[2]

$$\begin{aligned}\hat{H}\psi_{i,\vec{k}}(\vec{r}) &= \epsilon_{i,\vec{k}}\psi_{i,\vec{k}}(\vec{r}) \\ \hat{H}e^{i\vec{k}\cdot\vec{r}}u_{i,\vec{k}}(\vec{r}) &= \epsilon_{i,\vec{k}}e^{i\vec{k}\cdot\vec{r}}u_{i,\vec{k}}(\vec{r})\end{aligned}\quad (26)$$

Multiplying the second line in Eq. 26 by  $e^{-i\vec{k}\cdot\vec{r}}$ , we have

$$\begin{aligned}e^{-i\vec{k}\cdot\vec{r}}\hat{H}e^{i\vec{k}\cdot\vec{r}}u_{i,\vec{k}}(\vec{r}) &= \epsilon_{i,\vec{k}}u_{i,\vec{k}}(\vec{r}) \\ \hat{H}_{\vec{k}}u_{i,\vec{k}}(\vec{r}) &= \epsilon_{i,\vec{k}}u_{i,\vec{k}}(\vec{r})\end{aligned}\quad (27)$$

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.

### III. 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-von Karman boundary conditions with a period of  $N$  atoms.

We are going to employ what is known as the *tight-binding model* where the variational wavefunction is an expansion in terms of atomic orbitals located on different atoms. The character of orbitals to be included in the expansion depends on the species included. For a carbon chain, for instance, it is usually sufficient to take the  $2s$  and the  $2p$  orbitals since the  $1s$  orbitals do not contribute to bonding. However, here we shall only consider the  $1s$  orbitals to illustrate the idea.

Let us go through the steps of the calculations one by one :

1. Start with the atomic orbitals centered on each atom :

$$\phi_l(x - na) \quad (28)$$

where  $l$  is the orbital label (in three dimensions,  $l = 2s, 2p_x, 2p_y, 2p_z$  etc.),  $x$  is the coordinate variable,  $na$  is the location of the  $n$ th atom and  $a$  is the lattice constant.

2. Construct  $k$ -dependent functions from the orbitals in Eq. 28 to be used as a basis.

$$\chi_{kl}(x) = \frac{1}{\sqrt{N}} \sum_n e^{ikna} \phi_l(x - na) \quad (29)$$

where  $N$  is the number of atoms in the Born-von Karman cell. The normalization constant should be self-explanatory.

3. Check to make sure that  $\chi_{kl}$  obey Bloch's theorem.

$$\chi_{kl}(x + n'a) = \frac{1}{\sqrt{N}} \sum_n e^{ikna} \phi_l(x - (n - n')a) \quad (30)$$

Define a new variable for the summation  $n'' = n - n'$  and substitute into Eq. 30.

$$\begin{aligned} \chi_{kl}(x + n'a) &= \frac{1}{\sqrt{N}} \sum_{n''} e^{ik(n'+n'')a} \phi_l(x - n''a) \\ &= \frac{1}{\sqrt{N}} e^{ikn'a} \sum_{n''} e^{ikn''a} \phi_l(x - n''a) \\ &= e^{ikn'a} \chi_{kl}(x). \end{aligned} \quad (31)$$

4. Construct the variational wavefunction — do not forget that  $k$  is a good quantum number for this wavefunction.

$$\Psi_k(x) = \sum_l c_{kl} \chi_{kl}(x) \sum_{ln} c_{kl} e^{ikna} \phi_l(x - na) \quad (32)$$

5. Construct the variational equations to find the coefficients

$$\sum_{lm} \left[ \langle \chi_{km} | \hat{H} | \chi_{kl} \rangle - \varepsilon_k \langle \chi_{km} | \chi_{kl} \rangle \right] c_{kl} = 0 \quad (33)$$

6. Look at the overlap term in Eq. 33. Substitute Eq. 29 :

$$\begin{aligned} \langle \chi_{kl} | \chi_{km} \rangle &= \frac{1}{N} \sum_{nn'} e^{ik(n-n')a} \langle \phi_l(na) | \phi_m(n'a) \rangle \\ &= \frac{1}{N} \sum_{nn'} e^{ik(n-n')a} \langle \phi_l(0) | \phi_m(na - n'a) \rangle \end{aligned} \quad (34)$$

where  $|\phi_m(na)\rangle$  refers to an orbital centered on the  $n$ th atom located at  $x = na$ . Now let  $n'' = n - n'$

$$\begin{aligned} \langle \chi_{kl} | \chi_{km} \rangle &= \frac{1}{N} \sum_{n'n''} e^{ikn''a} \langle \phi_l(0) | \phi_m(n''a) \rangle \\ &= \sum_n e^{ikna} \langle \phi_l(0) | \phi_m(na) \rangle \end{aligned} \quad (35)$$

Eq. 35 is the most general expression for a general set of orbitals. In this simple example, we only consider one kind of orbital, namely  $l = 1s$  and as a result the sum over  $l$  in Eq. 35 reduces to a single term.

Next, we assume that the overlap between the orbitals on neighboring atoms is so small that the inner product of orbitals in Eq. 35 is nonzero only when  $n = 0$ , i.e. when both orbitals are on the same atom. This is not necessary but makes our job easier. Thus,

$$\langle \phi(0) | \phi(na) \rangle = \delta_{n,0} \quad (36)$$

7. Similarly for the matrix element of the Hamiltonian,

$$\langle \chi_{kl} | \hat{H} | \chi_{km} \rangle = \sum_n e^{ikna} \langle \phi(0) | \hat{H} | \phi(na) \rangle \quad (37)$$

This time, we assume that the Hamiltonian matrix element does not only remain on-site but extends to neighbors. This amounts to

$$\langle \phi(0) | \hat{H} | \phi(na) \rangle = \varepsilon_s \delta_{n,0} + t \delta_{n,\pm 1}. \quad (38)$$

This particular approximation is referred as the *nearest-neighbor approximation*.

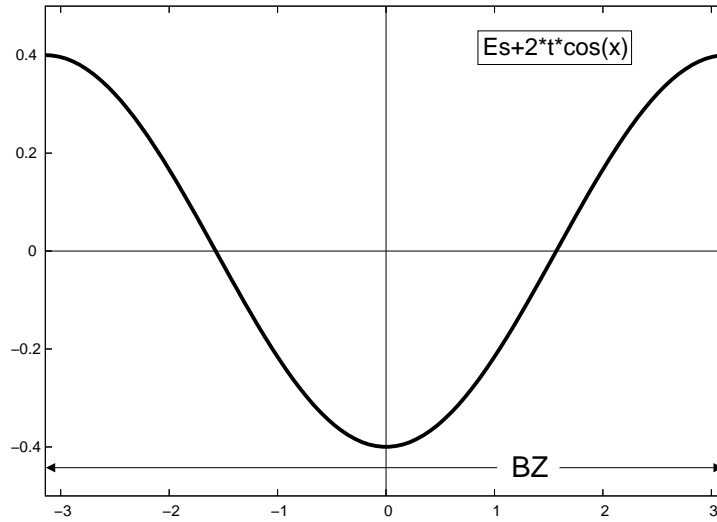
8. Substitute Eq. 36 and Eq. 38 into Eq. 33 to get

$$\sum_n e^{ikna} (\varepsilon_s \delta_{n,0} + t \delta_{n,\pm 1}) = \varepsilon_k \sum_n e^{ikna} \delta_{n,0}. \quad (39)$$

Note that even though the original equation started out as a matrix equation to determine the coefficients, the simplicity of the single-orbital assumption turned it into a single equation.

9. Finally, we obtain the following analytic equation for the energy eigenvalues

$$\varepsilon_k = \varepsilon_s + 2t \cos ka \quad (40)$$



[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 Kaxiras

# Lecture III : The Many-Body Hamiltonian and the Functional Derivative

## I. THE MANY-BODY HAMILTONIAN

Many of the properties of atoms, molecules and solids may be understood by determining the eigenfunctions of the *many-body* Hamiltonian,

$$\hat{H} = \hat{T} + \hat{V}. \quad (1)$$

For the relevant energy scale of such systems, we are interested in the energy contributions due to two kinds of particles : electrons and nuclei. Three types of interaction between these components thus contribute to the potential term  $\hat{V}$  :

- *Nucleus-nucleus* : Nuclei are taken as classical particles that interact via the Coulombic forces. This contribution is rather difficult to calculate for a solid but there is convenient way of handling this called the *Ewald sum*.
- *Nucleus-electron interaction* : This is a Coulombic interaction involving one electron at a time. Electrons are considered as quantum mechanical particles and thus this interaction behaves like an external single-body potential acting on the electrons. In practice, this part is taken care of by convenient *pseudopotentials* replacing the true potentials of the nuclei.
- *Electron-electron interaction* : This is again a Coulombic interaction involving pairs of electrons. This part of the Hamiltonian is what makes the many-body problem so hard, in most cases impossible, to tackle. Almost all electronic structure calculation methods resort to approximations which simplify the electron-electron interaction. The quality of the particular calculation used for a system depends on how well this approximation is chosen.

In theory the kinetic energy term is the sum of the kinetic energies of electrons and nuclei. However, because the ions are several tens of thousands of times heavier than the electrons, their contribution to the kinetic energy is usually neglected. This is called the *Born-Oppenheimer approximation*. Another way of stating this approximation is that the time scale of the motion of electrons is much larger than the nuclei and thus the response time of the electrons to any change in the positions of the nuclei is considered immediate.

In light of the Born-Oppenheimer approximation, only the electrons have a quantum mechanical identity in the problem at hand. Thus apart from a constant term brought around by the Coulombic interaction between nuclei, the eigenfunctions of the Hamiltonian depend only on electronic variables.

Before writing down the many-body Hamiltonian, let's remind ourselves of the case of the single electron in a Coulombic potential, i.e. the case of the hydrogen atom. The Hamiltonian in this case assumes the simple form,

$$\hat{H}_s = -\frac{\hbar^2}{2m_e} \nabla_{\vec{r}}^2 - \frac{e^2}{4\pi\epsilon_0} \frac{1}{r} \quad (2)$$

where  $m_e$  is the electron mass,  $\epsilon_0$  is the dielectric constant of vacuum and  $r = |\vec{r}|$  is the length of the position vector, with origin at the nucleus. The eigenfunctions of the Schrödinger equation

$$\hat{H}_s \psi(\vec{r}) = E \psi(\vec{r}) \quad (3)$$

have the well-known separable form

$$\psi(\vec{r})_{nlm} = \psi_{nl}(r) Y_{lm}(\theta, \phi) \quad (4)$$

where the radial part is given by Laguerre polynomials and the angular part (for any spherically symmetric potential) by spherical harmonics. Each solution is designated by three quantum numbers  $n$ ,  $l$  and  $m$ . We shall come back to the single-electron atom problem when we are dealing with pseudopotentials.

As soon as a second electron is added to the above system, as in the Helium atom, a new term emerges in the Hamiltonian which is the interaction term between electrons. We thus have

$$\hat{H}(\vec{r}_1, \vec{r}_2) = -\frac{\hbar^2}{2m_e} (\nabla_{\vec{r}_1}^2 + \nabla_{\vec{r}_2}^2) - \frac{Ze^2}{4\pi\epsilon_0} \left( \frac{1}{r_1} + \frac{1}{r_2} \right) + \frac{e^2}{4\pi\epsilon_0} \frac{1}{|\vec{r}_1 - \vec{r}_2|} \quad (5)$$

where we have a different position variable per electron,  $\vec{r}_i$  and  $Z = 2$ . The last term in Eq. 5 is the crucial ingredient which introduces complications into the problem by coupling the electronic variables. If this term were absent each of



the electrons would move independently as if the other one did not exist, the wavefunction would be separable in  $\vec{r}_1$  and  $\vec{r}_2$  and the energy spectrum would be identical to that of the single-particle problem, each level occupied by an electron obeying the Pauli principle. Because of the interaction term, however, the dependence of the wavefunction on  $\vec{r}_1$  and  $\vec{r}_2$  is no longer a product of functions, each depending on a single particle but rather has a complicated dependence on both variables simultaneously. The full wavefunction of course contains also the spin coordinates must be antisymmetric under the exchange of variables :

$$\psi(\vec{r}_1\sigma_1, \vec{r}_2\sigma_2) = -\psi(\vec{r}_2\sigma_2, \vec{r}_1\sigma_1). \quad (6)$$

In the Hamilton formalism the spin part and the spatial part of the wavefunctions may be separated and the spin degrees of freedom are added externally in agreement with the Pauli principle. We shall thus be concerned (at least initially) with the spatial part in what follows.

Next, let's generalize this to a system with  $N_e$  electrons and  $N_n$  nuclei.

$$\begin{aligned} \hat{H} &= \hat{T}_e + \hat{V}_{ne} + \hat{V}_{ee} + \hat{V}_{nn} \\ &= -\frac{\hbar^2}{2m_e} \sum_i \nabla_i^2 + \frac{e^2}{4\pi\epsilon_0} \left[ -\sum_i \sum_I \frac{Z_I}{|\vec{r}_i - \vec{R}_I|} + \frac{1}{2} \sum_i \sum_{j \neq i} \frac{1}{|\vec{r}_i - \vec{r}_j|} + \frac{1}{2} \sum_I \sum_{J \neq I} \frac{Z_I Z_J}{|\vec{R}_I - \vec{R}_J|} \right] \end{aligned} \quad (7)$$

where indices  $i$  and  $j$  run over the electronic degrees of freedom and  $I$  and  $J$  run over nuclei. The many-body Hamiltonian in Eq. 7 is a straightforward extension of the Helium Hamiltonian of Eq. 5. However, a few points should be paid attention to during this transition :

- The nuclei do not necessarily have the same  $Z$ .
- All the Coulombic interactions have double sums running over different components.
- In the inner sum of the last two terms, the  $i = j$  ( $I = J$ ) terms should be excluded so that electrons (nuclei) do not interact with themselves. This is not necessary for the electron-nuclei interaction (second term in Eq. 7).
- The factor of  $\frac{1}{2}$  in the third and fourth terms (electron-electron and nuclei-nuclei interaction) has been included in order to avoid double counting of terms. Again, this is not needed for the second term as the two entities are different.

Although the many-body Hamiltonian in Eq. 7 seems very straightforward, its exact solution is an impossible task for systems having more than a few electrons. Several methods have been devised to find accurate approximations. There are two broad classes of methods, *wavefunction-based* and *density-based* and each of these classes are further subdivided into different approaches

1. *Wavefunction-based methods* : An explicit form for the wavefunction is written down and observables are calculated using this wavefunction.
  - (a) Perturbational : Moller-Plesset, diagrammatic methods
  - (b) Variational : Hartree-Fock, configuration interaction (CI)
2. *Density-based methods* : The focus is shifted from the wavefunction to the electronic density. The wavefunction is not written explicitly. Examples are Thomas-Fermi approximation and density-functional theory.

A few final points remain to be mentioned before we move on to the next subsection. Several generalizations may be made to the many-body Hamiltonian in Eq. 7

- Other external single-body potentials may be added such as an electric or a magnetic field. Such a potential would be treated on equal footing with nucleus-electron interaction.
- Spin-orbit coupling may be included. In this case the wavefunctions are represented by *spinors*.

## II. ATOMIC UNITS

While writing the many-body Hamiltonian in Eq. 7, we have employed SI units. As a result of this, each term includes certain cumbersome combinations of fundamental constants  $\hbar$ ,  $m$ ,  $e$  and  $\epsilon_0$ . In addition to making the expression seem busy and hard to deal with, the SI units of energy(J) and length(m) are inappropriate when we are dealing with atomic systems. A more natural unit system to employ for this problem is the *atomic units*. In atomic units, the fundamental

energy and length is measured in terms of Hartree and the Bohr radius respectively. A Hartree is the ground state energy of the electron in the hydrogen atom and the Bohr radius is its average ground state radius :

$$a_0 = \frac{4\pi\epsilon_0\hbar^2}{m_e e^2}$$

$$E_h = \frac{\hbar^2}{m_e a_0^2} = \frac{e^2}{4\pi\epsilon_0} \quad (8)$$

In order to see how atomic units simplify the electronic Hamiltonian, let's start with Eq. 2 and rescale all the coordinates by  $a_0$ , i.e.  $\vec{r} \rightarrow \tilde{\vec{r}} = \vec{r}/a_0$ . Note that this has the effect of rescaling of the Laplacian operator,  $\nabla_{\vec{r}}^2 \rightarrow \nabla_{\tilde{\vec{r}}}^2/a_0^2$ . The Hamiltonian then becomes

$$\hat{H} = -\frac{\hbar^2}{2m_e a_0^2} \nabla_{\tilde{\vec{r}}}^2 - \frac{e^2}{4\pi\epsilon_0} \frac{1}{a_0} \frac{1}{\tilde{r}} = -\underbrace{\frac{\hbar^2}{m_e a_0^2} \frac{1}{2} \nabla_{\tilde{\vec{r}}}^2}_{E_H} - \underbrace{\frac{e^2 m_e}{4\pi\epsilon_0 \hbar^2} \frac{\hbar^2}{m_e a_0}}_{1/a_0} \frac{1}{\tilde{r}} = E_H \left( -\frac{1}{2} \nabla_{\tilde{\vec{r}}}^2 - \frac{1}{\tilde{r}} \right) \quad (9)$$

Now we rescale the Hamiltonian by  $E_H$ , i.e.  $\hat{H} \rightarrow \hat{H}/E_H$ . But because the choice of symbols are arbitrary, we are allowed to drop the  $\sim$ . We thus have the following simplified Hamiltonian

$$\hat{H} = -\frac{1}{2} \nabla_{\tilde{\vec{r}}}^2 - \frac{1}{\tilde{r}} \quad (10)$$

where the lengths are measured in Bohr ( $1 \text{ B} = 0.529 \text{ \AA}$ ) and energies in Hartree ( $1 \text{ H} = 27.211 \text{ eV}$ ). The rescaling operation outlined above is equivalent to setting  $\hbar = m_e = e = 1/4\pi\epsilon_0 = 1$ . Repeating the same procedure for the many-body Hamiltonian yields

$$\hat{H} = -\frac{1}{2} \sum_i^{N_e} \nabla_i^2 - \sum_i^{N_e} \sum_I^{N_n} \frac{Z_I}{|\vec{r}_i - \vec{R}_I|} + \frac{1}{2} \sum_i^{N_e} \sum_{j \neq i}^{N_e} \frac{1}{|\vec{r}_i - \vec{r}_j|} + \frac{1}{2} \sum_I^{N_n} \sum_{J \neq I}^{N_n} \frac{Z_I Z_J}{|\vec{R}_I - \vec{R}_J|}. \quad (11)$$

### III. THE FUNCTIONAL DERIVATIVE

The reason why the word *functional* appears in density functional theory is that the energy is written as a *functional* of the electronic density. While a function maps one or several variables to a number, a functional maps a *function* to a number :

$$f(x) : x \rightarrow y \quad \text{function}$$

$$F[f(x)] : f(x) \rightarrow y \quad \text{functional} \quad (12)$$

As seen in Eq. 12, arguments of functionals are written in square brackets to distinguish them from regular functions. The perimeter and area of a closed curve parametrized by an angle are familiar examples of functionals.

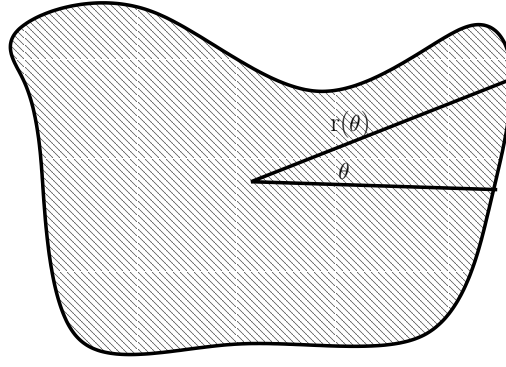
$$P[r] = \int_0^{2\pi} r(\theta) d\theta$$

$$A[r] = \frac{1}{2} \int_0^{2\pi} r^2(\theta) d\theta \quad (13)$$

As for regular functions, the concept of derivative also exists for functionals. When a small change is affected in the form of the function in the argument, the infinitesimal change in the value of the functional is measured by the *functional derivative*. The functional derivative should be the same for any arbitrary small change. The functional derivative is defined by

$$F[f(x) + \delta f(x)] - F[f(x)] = \int dx \frac{\delta F[f(x)]}{\delta f(x)} \delta f(x) \quad (14)$$

Taking for example the area of a parametrized curve defined in Eq. 13, we evaluate its functional derivative with respect to  $r$  using the definition in Eq. 14



$$\begin{aligned}
 A[r + \delta r] - A[r] &= \int \frac{1}{2}((r + \delta r)^2 - r^2)d\theta = \int \frac{1}{2}(r^2 + 2r\delta r - r^2)d\theta \\
 &= \int r \delta r d\theta \Rightarrow \frac{\delta A[r]}{\delta r} = r
 \end{aligned} \tag{15}$$

Functionals that are written as an integral of a simple function of their argument (such as  $P(\theta)$  and  $A(\theta)$  in Eq. 13) are referred to as *local functionals* since they depend only on the knowledge of the function at a given point at a time. In particular the most common functionals we shall encounter in the course of this lecture are those that are written in the form of a definite integral

$$F[f(x)] = \int_a^b dx a(f(x)) \tag{16}$$

Given this form, a general expression for their functional derivatives can be derived as

$$\int dx [a(f(x) + \delta f(x)) - a(f(x))] = \int dx \left[ a(f) + \frac{da}{df} \delta a(x) - a(f) \right] = \int dx \frac{da}{df} \delta a(x) \Rightarrow \frac{\delta a}{\delta f} = a'(f(x)) \tag{17}$$

In simple terms, if you have a local functional,  $F[f(x)]$  in the form of a definite integral, you just take the derivative of the integrand with respect to the function  $f$  and get rid of the integral sign.

In addition to local functionals, there also exist functionals which involve derivatives of the argument as well. In fact, we shall encounter such functionals while talking about the generalized gradient approximation to electron exchange and correlation.

Assuming that all derivatives of the function  $a$  vanishes at the limits of the integral, we can derive the functional derivative of a *semilocal functional*, which depends only on the derivatives only up to the first order as

$$\begin{aligned}
 F[f, f'] &= \int dx [a(f(x) + \delta f(x), f'(x) + \delta f'(x)) - a(f(x), f'(x))] \\
 &= \int dx \left[ a(f(x), f'(x)) + \frac{\partial a}{\partial f} \delta f(x) + \frac{\partial a}{\partial f'} \delta f'(x) \right] - a(f(x), f'(x)) \\
 &= \int dx \frac{\partial a}{\partial f} \delta f(x) + \int dx \frac{d}{dx} \frac{\partial a}{\partial f'} \delta f(x) - \int dx \frac{d}{dx} \frac{\partial a}{\partial f'} \delta f(x)
 \end{aligned} \tag{18}$$

In the last line of Eq. 18, the second term cancels due to the assumption of vanishing derivatives. The Second and third terms combine to give the functional derivative as

$$\frac{\delta F[f, f']}{\delta f} = \frac{\partial a}{\partial f} - \frac{d}{dx} \frac{\partial a}{\partial f'} \tag{19}$$

It is straightforward to generalize this formula to a nonlocal functional

$$F[f, f', f'', f''', \dots] = \int a(f, f', f'', f''', \dots) dx \tag{20}$$

that depends also on higher order derivatives of the function  $f$  as

$$\frac{\delta F[f, f', f'', f''', \dots]}{\delta f} = a - \frac{d}{dx} \frac{\partial a}{\partial f'} + \frac{d^2}{dx^2} \frac{\partial a}{\partial f''} - \frac{d^3}{dx^3} \frac{\partial a}{\partial f'''} + \dots \tag{21}$$

## Lecture IV : The Hartree-Fock method

### I. THE HARTREE METHOD

We have seen in the previous lecture that the many-body Hamiltonian for an electronic system may be written in atomic units as

$$\hat{H} = -\frac{1}{2} \sum_i^{N_e} \nabla_i^2 - \sum_i^{N_e} \sum_I^{N_n} \frac{Z_I}{|\vec{r}_i - \vec{R}_I|} + \frac{1}{2} \sum_i^{N_e} \sum_{j \neq i}^{N_e} \frac{1}{|\vec{r}_i - \vec{r}_j|} + \frac{1}{2} \sum_I^{N_n} \sum_{J \neq I}^{N_n} \frac{Z_I Z_J}{|\vec{R}_I - \vec{R}_J|}. \quad (1)$$

For the time being we take the final term to be a constant and we only consider the electronic Hamiltonian, which is just

$$\hat{H}_e = -\frac{1}{2} \sum_i^{N_e} \nabla_i^2 - \sum_i^{N_e} \sum_I^{N_n} \frac{Z_I}{|\vec{r}_i - \vec{R}_I|} + \frac{1}{2} \sum_i^{N_e} \sum_{j \neq i}^{N_e} \frac{1}{|\vec{r}_i - \vec{r}_j|} \quad (2)$$

In Eq. 2 the kinetic energy term and the nucleus-electron interaction term are sums of single-particle operators, each of which act on a single electronic coordinate. The electron-electron interaction term on the other hand is a pair interaction and acts on pairs of electrons. To facilitate the upcoming math, let's make the following definition

$$\hat{H}_e = \sum_i \hat{h}_1(\vec{x}_i) + \frac{1}{2} \sum_{i \neq j} \hat{h}_2(\vec{x}_i, \vec{x}_j). \quad (3)$$

where  $\vec{x}_i$  is now a generalized coordinate that includes spatial as well as spin degrees of freedom.

The Hartree-Fock method is a variational, wavefunction-based approach. Although it is a many-body technique, the approach followed is that of a single-particle picture, i.e. the electrons are considered as occupying single-particle orbitals making up the wavefunction. Each electron feels the presence of the other electrons indirectly through an effective potential. Each orbital, thus, is affected by the presence of electrons in other orbitals.

The starting point of the Hartree-Fock method is to write a variational wavefunction, which is built from these single-particle orbitals. Once we make a suitable *ansatz* to the wavefunction, all that is left is the application of the variational principle as described in Lecture 1. The simplest wavefunction that can be formed from these orbitals is their direct product

$$\Phi(\vec{x}_1, \dots, \vec{x}_N) = \phi_1(\vec{x}_1) \phi_2(\vec{x}_2) \dots \phi_N(\vec{x}_N). \quad (4)$$

This is the *Hartree* approximation and it is a straightforward task to calculate the variational lowest energy from Eq. 4. However, the Hartree wavefunction has a very important shortcoming, which is that it fails to satisfy antisymmetry, which states that a fermion wavefunction changes sign under odd permutations of the electronic variables. The permutation operator is defined by its action on the wavefunction

$$\hat{P}_{ij} \Phi(\vec{x}_1, \dots, \vec{x}_i, \dots, \vec{x}_j, \dots, \vec{x}_N) = \Phi(\vec{x}_1, \dots, \vec{x}_j, \dots, \vec{x}_i, \dots, \vec{x}_N) = -\Phi(\vec{x}_1, \dots, \vec{x}_i, \dots, \vec{x}_j, \dots, \vec{x}_N) \quad (5)$$

If an odd number of such permutation operators are applied to the wavefunction, it picks up a minus sign while no change in sign occurs under an even number of permutations. In order to satisfy the antisymmetry condition, a more sophisticated form than that of the Hartree wavefunction is needed.

### II. THE SLATER DETERMINANT

If, for example, we have a two-electron system with orbitals  $\phi_1(\vec{x}_1)$  and  $\phi_2(\vec{x}_2)$ , the following variational wavefunction satisfies the antisymmetry condition, at the same time preserving the single-particle picture

$$\Phi(\vec{x}_1, \vec{x}_2) = c [\phi_1(\vec{x}_1) \phi_2(\vec{x}_2) - \phi_1(\vec{x}_2) \phi_2(\vec{x}_1)] \quad (6)$$

where  $c$  is the normalization constant. For three electrons, the equivalent antisymmetrized wavefunction would be

$$\begin{aligned} \Phi(\vec{x}_1, \vec{x}_2, \vec{x}_3) = c [ & \phi_1(\vec{x}_1) \phi_2(\vec{x}_2) \phi_3(\vec{x}_3) - \phi_1(\vec{x}_1) \phi_2(\vec{x}_3) \phi_3(\vec{x}_2) + \phi_1(\vec{x}_3) \phi_2(\vec{x}_1) \phi_3(\vec{x}_2) \\ & - \phi_1(\vec{x}_2) \phi_2(\vec{x}_1) \phi_3(\vec{x}_3) + \phi_1(\vec{x}_3) \phi_2(\vec{x}_2) \phi_3(\vec{x}_1) - \phi_1(\vec{x}_2) \phi_2(\vec{x}_3) \phi_3(\vec{x}_1) ]. \end{aligned} \quad (7)$$

Upon closer inspection, we notice that the same permutations of orbitals with matching signs are obtained by the following determinant

$$\Phi = c \begin{vmatrix} \phi_1(\vec{x}_1) & \phi_2(\vec{x}_1) & \phi_3(\vec{x}_1) \\ \phi_1(\vec{x}_2) & \phi_2(\vec{x}_2) & \phi_3(\vec{x}_2) \\ \phi_1(\vec{x}_3) & \phi_2(\vec{x}_3) & \phi_3(\vec{x}_3) \end{vmatrix} \quad (8)$$

Generalizing this to an  $N$ -electron system where the orbitals are taken to satisfy orthonormality, we have

$$\Phi = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(\vec{x}_1) & \phi_2(\vec{x}_1) & \cdots & \phi_N(\vec{x}_1) \\ \phi_1(\vec{x}_2) & \phi_2(\vec{x}_2) & \cdots & \phi_N(\vec{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(\vec{x}_N) & \phi_2(\vec{x}_N) & \cdots & \phi_N(\vec{x}_N) \end{vmatrix} \quad (9)$$

where the factor in front ensures normalization. For an arbitrary number of electrons the wavefunction form in Eq. 9 can be shown to satisfy the desired antisymmetry condition. The determinant, referred to as a *Slater determinant* in literature, has  $N!$  terms each multiplied by -1 or 1 depending on the parity of the permutation. Each term has each orbital  $\phi_i$  only once and each of the arguments  $\vec{x}_i$  only once. Thus, each term may be written as follows :

$$(-1)^{P(i_1, i_2, \dots, i_N)} \phi_{i_1}(\vec{x}_1) \phi_{i_2}(\vec{x}_2) \cdots \phi_{i_N}(\vec{x}_N) \quad (10)$$

where the indices  $i_1, i_2, \dots$  take values between 1 and  $N$  and the exponent of -1 in front refers to the order of appearance of the orbital indices in the term. The term picks up a -1 in front if the corresponding permutation is odd and +1 if it is even. For ease of notation, we replace  $P(i_1, i_2, \dots, i_N)$  by the shorthand notation  $P(i)$ , where  $i$  now refers to a particular arrangement (or sequence) of the  $N$  indices. The Slater determinant may then be written as

$$\sum_i^{N!} (-1)^{P(i)} \phi_{i_1}(\vec{x}_1) \phi_{i_2}(\vec{x}_2) \cdots \phi_{i_N}(\vec{x}_N) \quad (11)$$

where the sum  $i$  runs over the  $N!$  terms. Nothing has been said so far about the form of the orbitals  $\phi_i(\vec{x}_j)$  and they are left to be found as a result of the minimization procedure associated by the variation. In order to achieve that we now calculate the expectation value of the Hamiltonian for this variational wavefunction

$$E_H = \langle \Phi | \hat{H}_e | \Phi \rangle. \quad (12)$$

### III. ELECTRONIC ENERGY FOR THE SLATER DETERMINANT

First, we tackle the single-particle terms in Eq. 2.

$$\langle \Phi | \sum_n \hat{h}_1(\vec{x}_n) | \Phi \rangle = \frac{1}{N!} \sum_n \sum_{i,j} (-1)^{P(i)} (-1)^{P(j)} \langle \phi_{j_1}(\vec{x}_1) \phi_{j_2}(\vec{x}_2) \cdots \phi_{j_N}(\vec{x}_N) | \hat{h}_1(\vec{x}_n) | \phi_{i_1}(\vec{x}_1) \phi_{i_2}(\vec{x}_2) \cdots \phi_{i_N}(\vec{x}_N) \rangle \quad (13)$$

Because each orbital making up the Slater determinant depends only on a single coordinate, we can pair up those orbitals that have the same argument and separate them into individual inner products, except for the one that has the same argument as the operator :

$$\begin{aligned} \langle \Phi | \sum_n \hat{h}_1(\vec{x}_n) | \Phi \rangle &= \frac{1}{N!} \sum_n \sum_{i,j} (-1)^{P(i)} (-1)^{P(j)} \times \\ &\quad \langle \phi_{j_1}(\vec{x}_1) | \phi_{i_1}(\vec{x}_1) \rangle \cdots \langle \phi_{j_{n-1}}(\vec{x}_{n-1}) | \phi_{i_{n-1}}(\vec{x}_{n-1}) \rangle \langle \phi_{j_n}(\vec{x}_n) | \hat{h}_1(\vec{x}_n) | \phi_{i_n}(\vec{x}_n) \rangle \times \\ &\quad \langle \phi_{j_{n+1}}(\vec{x}_{n+1}) | \phi_{i_{n+1}}(\vec{x}_{n+1}) \rangle \cdots \langle \phi_{j_N}(\vec{x}_N) | \phi_{i_N}(\vec{x}_N) \rangle \end{aligned} \quad (14)$$

Due to orthonormality of the chosen orbitals, all the inner products yield delta functions

$$\langle \Phi | \sum_n \hat{h}_1(\vec{x}_n) | \Phi \rangle = \frac{1}{N!} \sum_n \sum_{i,j} (-1)^{P(i)} (-1)^{P(j)} \delta_{j_1 i_1} \cdots \delta_{j_{n-1} i_{n-1}} \langle \phi_{j_n}(\vec{x}_n) | \hat{h}_1(\vec{x}_n) | \phi_{i_n}(\vec{x}_n) \rangle \delta_{j_{n+1} i_{n+1}} \cdots \delta_{j_N i_N} \quad (15)$$

Due to the Kronecker delta's all  $i_k$  are equal to  $j_k$  except for  $i_n$ . But because all terms appear exactly once in the products,  $i_n = j_n$  is also automatically satisfied. Thus, the sequence of indices labeled  $i$  is identical to that labeled  $j$ , making the permutations yield identical signs. We have, as a result,

$$\langle \Phi | \sum_n \hat{h}_1(\vec{x}_n) | \Phi \rangle = \frac{1}{N!} \sum_n \sum_i \langle \phi_{i_n}(\vec{x}_n) | \hat{h}_1(\vec{x}_n) | \phi_{i_n}(\vec{x}_n) \rangle \quad (16)$$

Now, for a given sequence labeled by  $i$  and for a fixed  $i_n$ , there are  $(N-1)!$  terms in the sum. The sum over sequence index  $i$  may then be reduced to the sum of a single index  $i_n$

$$\langle \Phi | \sum_n \hat{h}_1(\vec{x}_n) | \Phi \rangle = \frac{1}{N!} (N-1)! \sum_n \sum_{i_n} \langle \phi_{i_n}(\vec{x}_n) | \hat{h}_1(\vec{x}_n) | \phi_{i_n}(\vec{x}_n) \rangle \quad (17)$$

Because the expectation value is an integration over the variable  $\vec{x}_n$ , each term yields the same result and there are  $N$  such terms. Finally we replace the now arbitrary index  $i_n$  by a generic index  $i$ . The final expression is then

$$\langle \Phi | \sum_n \hat{h}_1(\vec{x}_n) | \Phi \rangle = \sum_i \langle \phi_i | \hat{h}_1 | \phi_i \rangle \quad (18)$$

Next, we deal with the more complicated case of the two-particle terms. The starting point is identical to the case of the single-body part of the Hamiltonian where orbitals sharing the same argument are paired up into inner products except for the ones that have the same argument as the operator. In the case of the two-body part of the Hamiltonian there are two such orbitals. Assuming without loss of generality that  $n < m$  we thus have

$$\langle \Phi | \frac{1}{2} \sum_{n \neq m} \hat{h}_2(\vec{x}_n, \vec{x}_m) | \Phi \rangle = \frac{1}{N!} \frac{1}{2} \sum_{n \neq m} \sum_{i,j} (-1)^{P(i)} (-1)^{P(j)} \delta_{j_1 i_1} \cdots \delta_{j_{n-1} i_{n-1}} \delta_{j_{n+1} i_{n+1}} \cdots \delta_{j_{m-1} i_{m-1}} \delta_{j_{m+1} i_{m+1}} \cdots \delta_{j_N i_N} \times \\ \langle \phi_{j_n}(\vec{x}_n) \phi_{j_m}(\vec{x}_m) | \hat{h}_2(\vec{x}_n, \vec{x}_m) | \phi_{i_n}(\vec{x}_n) \phi_{i_m}(\vec{x}_m) \rangle \quad (19)$$

This time, there are two cases regarding the values  $i_n$  and  $i_m$  can take :

- i.  $j_n = i_n$  and  $j_m = i_m$
- ii.  $j_n = i_m$  and  $j_m = i_n$ .

The first case is very similar to the situation in the treatment of the single-body term and causes the sequences  $i$  and  $j$  to be equal. For the second case however, the sequences differ by a single pair, where one has  $i_m$  and  $i_n$  interchanged. This term then picks up a minus sign. We may now reduce one of the sums over the sequences and obtain

$$\langle \Phi | \frac{1}{2} \sum_{n \neq m} \hat{h}_2(\vec{x}_n, \vec{x}_m) | \Phi \rangle = \frac{1}{N!} \frac{1}{2} \sum_{n \neq m} \sum_i^{N!} [\langle \phi_{i_n}(\vec{x}_n) \phi_{i_m}(\vec{x}_m) | \hat{h}_2(\vec{x}_n, \vec{x}_m) | \phi_{i_n}(\vec{x}_n) \phi_{i_m}(\vec{x}_m) \rangle - \\ \langle \phi_{i_m}(\vec{x}_n) \phi_{i_n}(\vec{x}_m) | \hat{h}_2(\vec{x}_n, \vec{x}_m) | \phi_{i_n}(\vec{x}_n) \phi_{i_m}(\vec{x}_m) \rangle] \quad (20)$$

By a similar argument to the one that leads to Eq. 18, we notice that for a fixed sequence  $i$  and for fixed  $i_n$  and  $i_m$ , there are  $(N-2)!$  terms in the sum. We do not divide by two because this factor is already included. We do however, make sure that  $i_n \neq i_m$  because each term shows up only once in the sum. Eq. 21 then becomes

$$\langle \Phi | \frac{1}{2} \sum_{n \neq m} \hat{h}_2(\vec{x}_n, \vec{x}_m) | \Phi \rangle = \frac{1}{N!} (N-2)! \frac{1}{2} \sum_{n \neq m} \sum_{i_n \neq i_m}^N [\langle \phi_{i_n}(\vec{x}_n) \phi_{i_m}(\vec{x}_m) | \hat{h}_2(\vec{x}_n, \vec{x}_m) | \phi_{i_n}(\vec{x}_n) \phi_{i_m}(\vec{x}_m) \rangle - \\ \langle \phi_{i_m}(\vec{x}_n) \phi_{i_n}(\vec{x}_m) | \hat{h}_2(\vec{x}_n, \vec{x}_m) | \phi_{i_n}(\vec{x}_n) \phi_{i_m}(\vec{x}_m) \rangle] \quad (21)$$

If  $n$  is not equal to  $m$ , for each  $n$  there are  $N-1$  different values for  $m$ . Thus the number of  $(n, m)$  pairs is  $N(N-1)$  and we may replace the sum over such pairs with this factor. Finally, we replace  $i_n$  and  $i_m$  by more conventional indices such as  $i$  and  $j$  and the two-body term reduces to

$$\langle \Phi | \frac{1}{2} \sum_{n \neq m} \hat{h}_2(\vec{x}_n, \vec{x}_m) | \Phi \rangle = \sum_{i,j} \frac{1}{2} \sum_{i,j}^N [\langle \phi_i \phi_j | \hat{h}_2 | \phi_i \phi_j \rangle - \langle \phi_j \phi_i | \hat{h}_2 | \phi_i \phi_j \rangle] \quad (22)$$

where we have once again removed the dependence on orbitals. Notice the important fact that the term  $i = j$  does not give any contribution to the sum because the two terms exactly cancel in that case. This is going to be a major difference between the Hartree-Fock method and density functional theory.

Putting together Eq. 18 and Eq. 22, the expectation value of the Hamiltonian for the Slater determinant is

$$\langle \Phi | \hat{H}_e | \Phi \rangle = \sum_i^N \langle \phi_i | \hat{h}_1 | \phi_i \rangle + \frac{1}{2} \sum_{i,j}^N \left[ \langle \phi_i \phi_j | \hat{h}_2 | \phi_i \phi_j \rangle - \langle \phi_j \phi_i | \hat{h}_2 | \phi_i \phi_j \rangle \right] \quad (23)$$

Next, we would like to apply the variational principle to the value in Eq. 23 to obtain the best possible estimate.

#### IV. THE HARTREE-FOCK EQUATIONS

The variational principle that we will apply here is rather different from the linear variation that we saw in the first lecture. There the form of our approximate wavefunction was written as an expansion over a collection of predetermined functions and we minimized the expectation value (at the same time obeying the normalization constraint) with respect to the coefficients of the basis functions. Here however we employ a much more general treatment where we minimize with respect to the basis functions themselves! Needless to say, this requires functional differentiation where any change affected in the expectation value in Eq. 23 due to an infinitesimal change in any of the orbitals  $\phi_k$  should be zero.

$$\phi_k \rightarrow \phi_k + \delta\phi_k \Rightarrow \delta\langle \Phi | \hat{H}_e | \Phi \rangle = 0 \quad (24)$$

In addition, we demand through Lagrange multipliers that the set of orbitals  $\phi_k$  remain orthogonal throughout the minimization process. The condition to be satisfied is then

$$\delta F \equiv \delta \left[ \langle \Phi | \hat{H}_e | \Phi \rangle - \sum_{i,j} \lambda_{ij} (\langle \phi_i | \phi_j \rangle - \delta_{ij}) \right] = 0 \quad (25)$$

where the number of the Lagrange multipliers is now  $N^2$  and we demand that the derivatives with respect to all the Lagrange multipliers vanish.

First, let's see how the single-body term changes under an infinitesimal change in one of the orbitals,  $\phi_k$ .

$$\delta\langle \Phi | \sum_i \hat{h}_1(\vec{x}_i) | \Phi \rangle = \langle \delta\phi_k | \hat{h}_1 | \phi_k \rangle + \langle \phi_k | \hat{h}_1 | \delta\phi_k \rangle = \langle \delta\phi_k | \hat{h}_1 | \phi_k \rangle + \langle \delta\phi_k | \hat{h}_1 | \phi_k \rangle^* \quad (26)$$

The last equality of Eq. 26 has been written making use of the hermiticity of the single-particle operator.

Similarly, the variation of the two-body term yields

$$\begin{aligned} \delta\langle \Phi | \frac{1}{2} \sum_{i \neq j} \hat{h}_2(\vec{x}_i, \vec{x}_j) | \Phi \rangle &= \delta\frac{1}{2} \sum_{i,j} \left\{ \langle \phi_i \phi_j | \hat{h}_2 | \phi_i \phi_j \rangle - \langle \phi_j \phi_i | \hat{h}_2 | \phi_i \phi_j \rangle \right\} \\ &= \frac{1}{2} \sum_i \left\{ \langle \phi_i \delta\phi_k | \hat{h}_2 | \phi_i \phi_k \rangle + \langle \phi_i \phi_k | \hat{h}_2 | \phi_i \delta\phi_k \rangle - \langle \delta\phi_k \phi_i | \hat{h}_2 | \phi_i \phi_k \rangle - \langle \phi_k \phi_i | \hat{h}_2 | \phi_i \delta\phi_k \rangle \right\} \\ &\quad + \frac{1}{2} \sum_j \left\{ \langle \delta\phi_k \phi_j | \hat{h}_2 | \phi_k \phi_j \rangle + \langle \phi_k \phi_j | \hat{h}_2 | \delta\phi_k \phi_j \rangle - \langle \phi_j \delta\phi_k | \hat{h}_2 | \phi_k \phi_j \rangle - \langle \phi_j \phi_k | \hat{h}_2 | \delta\phi_k \phi_j \rangle \right\} \end{aligned} \quad (27)$$

where it is understood that in the second line  $\hat{h}_2$  carries indices  $i$  and  $k$  and in the third line it carries indices  $j$  and  $k$ . We can simplify the equation in Eq. 27 by making use of the hermiticity of  $\hat{h}_2$  and the identity

$$\langle f_1 f_2 | \hat{A} | f_3 f_4 \rangle = \langle f_2 f_1 | \hat{A} | f_4 f_3 \rangle \quad (28)$$

which holds if *both* of the functions on either side of the two-body operator are interchanged simultaneously. Interchanging the order of orbitals on the third line yields the same terms as in the second line except for the summation indices  $i$  and  $j$  which are arbitrary and interchangeable. Having two identical sums, the factor of  $\frac{1}{2}$  cancels. Applying finally the Hermiticity, we reduce the two-body term to

$$\begin{aligned} \delta\langle \Phi | \frac{1}{2} \sum_{i \neq j} \hat{h}_2(\vec{x}_i, \vec{x}_j) | \Phi \rangle &= \sum_i \left[ \langle \phi_i \delta\phi_k | \hat{h}_2 | \phi_i \phi_k \rangle + \langle \phi_i \delta\phi_k | \hat{h}_2 | \phi_i \phi_k \rangle^* \right. \\ &\quad \left. - \langle \delta\phi_k \phi_i | \hat{h}_2 | \phi_i \phi_k \rangle - \langle \delta\phi_k \phi_i | \hat{h}_2 | \phi_i \phi_k \rangle^* \right] \end{aligned} \quad (29)$$

The functional derivative of the Lagrange multipliers proceeds in exactly the same way with the Hermiticity considered again at the end. Putting everything together the variation turns out to be

$$\begin{aligned} \delta F = & \langle \delta \phi_k | \hat{h}_1 | \phi_k \rangle + \langle \delta \phi_k | \hat{h}_1 | \phi_k \rangle^* + \sum_i [\langle \phi_i \delta \phi_k | \hat{h}_2 | \phi_i \phi_k \rangle + \langle \phi_i \delta \phi_k | \hat{h}_2 | \phi_i \phi_k \rangle^* \\ & - \langle \delta \phi_k \phi_i | \hat{h}_2 | \phi_i \phi_k \rangle - \langle \delta \phi_k \phi_i | \hat{h}_2 | \phi_i \phi_k \rangle^*] - \sum_i [\lambda_{ik} \langle \delta \phi_k | \phi_i \rangle^* + \lambda_{ki} \langle \delta \phi_k | \phi_i \rangle] \end{aligned} \quad (30)$$

For reasons which will become clear shortly, we shall evaluate the variation with respect to  $\phi_k^*$  and not  $\phi_k$ . (For a proof that variation with respect to  $\phi_k$  and  $\phi_k^*$  are equivalent, see the appendix at the end.) Written as an integral, terms in Eq. 30 would read, for instance,

$$\langle \delta \phi_k \phi_i | \hat{h}_2 | \phi_i \phi_k \rangle = \int \int \delta \phi_k^*(\vec{x}_1) \phi_i^*(\vec{x}_2) \hat{h}_2(\vec{x}_1, \vec{x}_2) \phi_i(\vec{x}_1) \phi_k(\vec{x}_2) d\vec{x}_1 d\vec{x}_2. \quad (31)$$

Moreover, since we are considering the variation with respect to only  $\delta \phi_k^*$ , we would completely disregard terms with  $\delta \phi$  in the ket. As we saw in the section on functional integrals, evaluating  $\frac{\delta F}{\delta \phi_k^*}$  basically amounts to getting rid of  $\delta \phi_k$  in the bras in Eq. 30 and also the integral over its argument. This leaves

$$\hat{h}_1 \phi_k(\vec{x}_1) + \sum_i \left\{ \int \phi_i^*(\vec{x}_2) \hat{h}_2[\phi_i(\vec{x}_2) \phi_k(\vec{x}_1)] d\vec{x}_2 - \int \phi_i^*(\vec{x}_2) \hat{h}_2[\phi_i(\vec{x}_1) \phi_k(\vec{x}_2)] d\vec{x}_2 \right\} = \sum_i \lambda_{ki} \phi_i(\vec{x}_1) \quad (32)$$

It is customary to write Eq. 32 using the orbital-dependent operators  $\hat{J}_i$  and  $\hat{K}_i$  as

$$\left[ \hat{h}_1 + \sum_i (\hat{J}_i - \hat{K}_i) \right] \phi_k = \sum_i \lambda_{ki} \phi_i \quad (33)$$

where the first and the second term are straightforward, single-body operators and the third term is an integral operator. This is now a set of interdependent single-particle eigenvalue equations. The operator  $\hat{J}$  corresponds to the classical interaction of an electron distributions given by  $|\phi_i|^2$  and  $|\phi_k|^2$  and is called the *direct term* while  $\hat{K}$ , called the *exchange term*, has no classical analogue and is a direct result of the antisymmetry property of the wavefunction. The *Fock operator* is defined as

$$\hat{F} = \hat{h}_1 + \sum_i (\hat{J}_i - \hat{K}_i) \quad (34)$$

and using this definition Eq. 33 takes the simple form

$$\hat{F} \phi_k = \sum_{i=1}^N \lambda_{ki} \phi_i \quad (35)$$

There are several different solutions to the equations in Eq. 35 each corresponding to a different set of  $\lambda_{ki}$ . We have the freedom to concentrate upon those  $\lambda_{ki}$  which satisfies

$$\lambda_{ki} = \delta_{ki} \epsilon_k \quad (36)$$

where  $\epsilon_k$  is essentially a new name for the Lagrange multipliers[1]. With this, Eq. 35 may be written as

$$\hat{F} \phi_k = \epsilon_k \phi_k. \quad (37)$$

In this form, Eq. 37 is a traditional eigenvalue equation. For each  $k$  there is an equivalent equation defining a system of Schrödinger-like, one-particle equations. Although it's tempting to interpret the eigenvalues  $\epsilon_k$  as the energy levels of an interacting system, this is in fact not justified because the single-electron picture is not correct. However, if interpreted correctly the Hartree-Fock eigenvalues do correspond to certain physical entities.

## V. TOTAL ENERGY AND KOOPMANS' THEOREM

The total Hartree-Fock energy,  $E_{HF}$ , of the system of  $N$  electrons discussed above may be written as

$$E_{HF} = \sum_i^N \langle \phi_i | \hat{h}_1 | \phi_i \rangle + \frac{1}{2} \sum_{i,j}^N [\langle \phi_i \phi_j | \hat{h}_2 | \phi_i \phi_j \rangle - \langle \phi_j \phi_i | \hat{h}_2 | \phi_i \phi_j \rangle] \quad (38)$$



making use of Eq. 18 and Eq. 22. On the other hand the eigenenergies,  $\epsilon_k$ , of the Hartree-Fock equations given in Eq. 37 may be summed to give

$$\sum_k \epsilon_k = \sum_i^N \langle \phi_i | \hat{h}_1 | \phi_i \rangle + \sum_{i,j}^N [\langle \phi_i \phi_j | \hat{h}_2 | \phi_i \phi_j \rangle - \langle \phi_j \phi_i | \hat{h}_2 | \phi_i \phi_j \rangle]. \quad (39)$$

Eq. 38 and Eq. 39 demonstrate that the total energy cannot be given as a simple sum over the eigenenergies of the Fock operator but are closely related by

$$E_{HF} = \sum_k \epsilon_k - \frac{1}{2} \sum_{i,j}^N [\langle \phi_i \phi_j | \hat{h}_2 | \phi_i \phi_j \rangle - \langle \phi_j \phi_i | \hat{h}_2 | \phi_i \phi_j \rangle]. \quad (40)$$

This “discrepancy” stems from the description of a true many-body system in terms of single-particle orbitals. Next, let’s compare the energy of two systems, one with  $N$  electrons and one with  $N - 1$ , where an electron has been removed from the  $n$ th level from the system with  $N$  electrons to obtain the system with  $N - 1$  electrons. The energy required to remove the electron is given by

$$\begin{aligned} \Delta E_N &\equiv E_{HF}(N - 1) - E_{HF}(N) \\ &= \sum_{i \neq n}^N \langle \phi_i | \hat{h}_1 | \phi_i \rangle + \frac{1}{2} \sum_{i,j \neq n}^N [\langle \phi_i \phi_j | \hat{h}_2 | \phi_i \phi_j \rangle - \langle \phi_j \phi_i | \hat{h}_2 | \phi_i \phi_j \rangle] \\ &\quad - \sum_i^N \langle \phi_i | \hat{h}_1 | \phi_i \rangle - \frac{1}{2} \sum_{i,j}^N [\langle \phi_i \phi_j | \hat{h}_2 | \phi_i \phi_j \rangle - \langle \phi_j \phi_i | \hat{h}_2 | \phi_i \phi_j \rangle] \\ &= -\langle \phi_n | \hat{h}_1 | \phi_n \rangle - \frac{1}{2} \sum_i^N [\langle \phi_i \phi_n | \hat{h}_2 | \phi_i \phi_n \rangle - \langle \phi_n \phi_i | \hat{h}_2 | \phi_i \phi_n \rangle] - \frac{1}{2} \sum_j^N [\langle \phi_n \phi_j | \hat{h}_2 | \phi_n \phi_j \rangle - \langle \phi_j \phi_n | \hat{h}_2 | \phi_n \phi_j \rangle]. \end{aligned} \quad (41)$$

which is obtained by evaluating the difference in a straightforward manner. In the final sum  $i$  can be substituted by  $j$  since the index is arbitrary. We then obtain

$$\Delta E_N = -\langle \phi_n | \hat{h}_1 | \phi_n \rangle - \frac{1}{2} \sum_i^N [\langle \phi_i \phi_n | \hat{h}_2 | \phi_i \phi_n \rangle - \langle \phi_n \phi_i | \hat{h}_2 | \phi_i \phi_n \rangle] - \frac{1}{2} \sum_i^N [\langle \phi_n \phi_i | \hat{h}_2 | \phi_n \phi_i \rangle - \langle \phi_i \phi_n | \hat{h}_2 | \phi_n \phi_i \rangle]. \quad (42)$$

In Eq. 42, it is permissible to exchange the indices  $i$  and  $n$  if done simultaneously on either side of the operator, we have

$$\Delta E_N = -\langle \phi_n | \hat{h}_1 | \phi_n \rangle - \frac{1}{2} \sum_i^N [\langle \phi_i \phi_n | \hat{h}_2 | \phi_i \phi_n \rangle - \langle \phi_n \phi_i | \hat{h}_2 | \phi_i \phi_n \rangle] - \frac{1}{2} \sum_i^N [\langle \phi_i \phi_n | \hat{h}_2 | \phi_i \phi_n \rangle - \langle \phi_n \phi_i | \hat{h}_2 | \phi_i \phi_n \rangle]$$

The two sums at the end are identical. Thus

$$\Delta E_N = -\langle \phi_n | \hat{h}_1 | \phi_n \rangle - \sum_i^N [\langle \phi_i \phi_n | \hat{h}_2 | \phi_i \phi_n \rangle - \langle \phi_n \phi_i | \hat{h}_2 | \phi_i \phi_n \rangle] \quad (43)$$

By comparing Eq. 43 and Eq. 35, we see that the right hand side is nothing but  $-\epsilon_n$

$$E_{HF}(N - 1) - E_{HF}(N) = -\epsilon_n \quad (44)$$

We thus arrive at *Koopmans’ theorem*, which states that each eigenvalue of the Fock operator gives the energy required to remove an electron from the corresponding single-electron state. Similarly the energy required to add an electron to orbital  $m$  can be proven to be given by

$$E_{HF}(N + 1) - E_{HF}(N) = \epsilon_m \quad (45)$$

In Eq. 44 and Eq. 45, level  $n$  is assumed to be occupied and level  $m$  is assumed to be empty. Therefore  $\epsilon_k$  may be interpreted as orbital energies in the approximation that the changes in the orbitals when the electron is removed or added is small.

## VI. HARTREE-FOCK-ROOTHAAN EQUATIONS

Solving the Fock equations given in Eq. 33 involves the evaluation of the orbitals at every single point in space, which is a rather daunting task. Instead, Roothaan proposed a limited variation instead where the orbitals are expanded in terms of a known basis set and the variation is conducted over the coefficients of the basis functions. This is slightly different from the usual variational principle that we discussed in Lecture 1 since there the approximate wavefunction was directly expanded in terms of the known basis without the extra step of the Slater determinant inbetween. In Roothaan's expansion, each orbital is written as

$$\phi_l(\vec{x}) = \sum_p^{N_b} c_{pl} \chi_p(\vec{x}) \quad (46)$$

where  $N_b$  is the number of basis sets used. Most of the time, the basis that we use is going to be a truncated set and  $N_b$  is going to be finite. When substituted into Eq. 25, this expansion yields

$$\begin{aligned} \hat{F} = & \sum_{i=1}^N \sum_{m,n=1}^{N_b} c_{mi}^* c_{ni} \langle \chi_m | \hat{h}_1 | \chi_n \rangle + \\ & \frac{1}{2} \sum_{i,j}^N \sum_{m,n,p,q}^{N_b} c_{mi}^* c_{pj} c_{nj}^* c_{qj} \left[ \langle \chi_m \chi_p | \hat{h}_2 | \chi_n \chi_q \rangle - \langle \chi_p \chi_m | \hat{h}_2 | \chi_n \chi_q \rangle \right] - \sum_i^N \epsilon_i \left[ \sum_{n,m}^{N_b} c_{mi}^* c_{ni} \langle \chi_m | \chi_n \rangle - 1 \right] \end{aligned} \quad (47)$$

where we once again make use of diagonal Lagrange multipliers. Since we conduct the variation over the coefficients rather than the orbitals, we have

$$\delta F = 0 \quad \Rightarrow \quad \frac{\partial F}{\partial c_{rk}^*} = 0 \quad (48)$$

The first term will yield a nonzero result only if  $m = r$  and  $k = i$ , which reduces to

$$\sum_n^{N_b} c_{nk} \langle \chi_r | \hat{h}_1 | \chi_n \rangle. \quad (49)$$

The second term brings two contributions, one for each of the  $c^*$  coefficients. By previous experience (and also by a quite straightforward yet tedious calculation) we foresee that these two contributions will be equal. Let's then look at the case where  $m = r$  and  $k = i$  and multiply by two to obtain the entire contribution from the two-body term

$$\sum_j^N \sum_{p,q}^{N_b} c_{pk} c_{nj}^* c_{qj} \left[ \langle \chi_r \chi_p | \hat{h}_2 | \chi_n \chi_q \rangle - \langle \chi_p \chi_r | \hat{h}_2 | \chi_n \chi_q \rangle \right] \quad (50)$$

For further use, let's exchange the summation indices  $p$  and  $n$  to obtain

$$\sum_j^N \sum_{p,q}^{N_b} c_{nk} c_{pj}^* c_{qj} \left[ \langle \chi_r \chi_n | \hat{h}_2 | \chi_p \chi_q \rangle - \langle \chi_n \chi_r | \hat{h}_2 | \chi_p \chi_q \rangle \right] \quad (51)$$

Finally the term with the Lagrange multipliers can be evaluated quite simply to result in

$$\epsilon_k \sum_n^{N_b} c_{nk} \langle \chi_r | \chi_n \rangle \quad (52)$$

The minimization then yields, putting everything together and collecting all terms under a single sum

$$\sum_n^{N_b} \left\{ \langle \chi_r | \hat{h}_1 | \chi_n \rangle + \sum_j^N \sum_{p,q}^{N_b} c_{pj}^* c_{qj} \left[ \langle \chi_k \chi_n | \hat{h}_2 | \chi_p \chi_q \rangle - \langle \chi_n \chi_k | \hat{h}_2 | \chi_p \chi_q \rangle \right] \right\} c_{nk} = \epsilon_k \sum_n^{N_b} \langle \chi_r | \chi_n \rangle c_{nk} \quad (53)$$

If now, we call the term in the large curly braces  $F_{rn}$  and recognize the inner product on the right-hand side as an element of the overlap matrix, namely  $S_{rn}$  then Eq. 53 turns into the following matrix equation

$$F \cdot C = \epsilon C \quad (54)$$

where  $C$  is the usual vector of coefficients.

Eq. 54, although very similar to the matrix equation that we obtained when dealing with the variational principle, differs in that the representation of the matrix  $F$  in any basis depends again on the coefficients that we are trying to calculate. This is again a result of the description of the many-body wavefunction as formed by noninteracting single-particle orbitals where the effect of the orbitals come out indirectly in the operator that defines the eigenvalue equation. A very similar problem will be encountered when we deal with DFT and standard methods of tackling with such so called *self-consistent field* problems will be illustrated.

## VII. SPIN PROPERTIES

Because the Slater determinant is a product of imagination, it is no guaranteed to satisfy all the properties of the exact wavefunction. Let us, for instance, look at the spin properties of the Slater determinant. We shall consider two operators :  $\hat{S}_z$  and  $\hat{S}^2$ . The total spin in any one direction is the sum of the spins of individual electrons

$$\hat{S}_x = \sum_i^N \hat{S}_{i,x} \quad \hat{S}_y = \sum_i^N \hat{S}_{i,y} \quad \text{and} \quad \hat{S}_z = \sum_i^N \hat{S}_{i,z} \quad (55)$$

and the total spin is given by

$$\hat{S}^2 = \hat{S}_x^2 + \hat{S}_y^2 + \hat{S}_z^2 \quad (56)$$

In order to find out about the spin properties of the approximate function, we need to explicitly specify the spin dependence of the individual orbitals. Let's assume that  $n$  of the orbitals have  $\alpha$  spin while  $m$  of them have  $\beta$  spin. Then the determinant may be written in shorthand notation as

$$\Phi(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N) = |\phi_1^+ \alpha, \phi_2^+ \alpha, \dots, \phi_n^+ \alpha, \phi_1^- \beta, \phi_2^- \beta \dots \phi_m^- \beta| \quad (57)$$

where we have allowed the position-dependent  $\phi^+$  and  $\phi^-$  orbitals to be different. If the function,  $\Phi$  in Eq. 57 is written down explicitly as a sum, we have

$$\Phi = \frac{1}{\sqrt{N!}} \sum_{i=1}^{N!} (-1)^{P(i)} \phi_{i_1}^+(\vec{r}_1) \alpha(1) \dots \phi_{i_n}^+(\vec{r}_n) \alpha(n), \phi_{i_{n+1}}^-(\vec{r}_{n+1}) \beta(n+1) \dots \phi_{i_N}^-(\vec{r}_N) \quad (58)$$

Using the  $\hat{S}_z$  operator in Eq. 55, the wavefunction in Eq. 58 is clearly seen to be an eigenstate of the  $\hat{S}_z$  operator with

$$\hat{S}_z \Phi = \frac{n-m}{2} \Phi \quad (59)$$

One can prove, however, through some straightforward (but somewhat lengthy) algebra that  $\Phi$  is an eigenfunction of  $S^2$  only if  $m = n$  and  $\phi^- = \phi^+$ . If we impose these constraints on the Slater determinant, this amounts to the *restricted Hartree-Fock method*. Alternatively, we may either ignore the problem and use *unrestricted Hartree-Fock* or resort to the *projected Hartree-Fock method* where the linear combination of more than one Slater determinant is taken as the approximate wavefunction, which is also an eigenfunction of the operator  $\hat{S}^2$ .

## VIII. APPENDIX

Consider a real function,  $f(z, z^*)$ , of the complex variable  $z$  and its complex conjugate  $z^*$ . The real and imaginary parts of the variable  $z$  can be written as linear combinations of  $z$  and  $z^*$  as

$$\begin{aligned} z_r &= \frac{z + z^*}{2} \\ z_i &= \frac{z - z^*}{2i}. \end{aligned} \quad (60)$$

If we consider now the derivatives with respect to  $z$  and make use of the chain rule, we obtain

$$\left. \frac{\partial f}{\partial z} \right|_{z^*} = \frac{\partial f}{\partial z_r} \frac{\partial z_r}{\partial z} + \frac{\partial f}{\partial z_i} \frac{\partial z_i}{\partial z} = \frac{1}{2} \left( \frac{\partial f}{\partial z_r} - i \frac{\partial f}{\partial z_i} \right) \quad (61)$$

where the last inequality has been written using Eq. 60. Similarly the derivative with respect to  $z^*$  yields

$$\left. \frac{\partial f}{\partial z^*} \right|_z = \frac{\partial f}{\partial z_r} \frac{\partial z_r}{\partial z^*} + \frac{\partial f}{\partial z_i} \frac{\partial z_i}{\partial z^*} = \frac{1}{2} \left( \frac{\partial f}{\partial z_r} + i \frac{\partial f}{\partial z_i} \right). \quad (62)$$

Since  $f$  is real, the derivatives  $\frac{\partial f}{\partial z_r}$  and  $\frac{\partial f}{\partial z_i}$  in Eq. 61 and Eq. 62 are also real. Equating both of these equations to zero at the minimum then yields the conditions

$$\frac{\partial f}{\partial z_r} = 0 \quad \text{and} \quad \frac{\partial f}{\partial z_i} = 0 \quad (63)$$

and are thus equivalent.

---

[1] One can prove that this it is possible to move between the diagonal form of the Lagrange multipliers and the more general non-diagonal form through a unitary transform

## I. THE DEFINITION OF DENSITY

As we have stressed several times in the lectures, solving the many-particle Schrödinger equation for the ground state wavefunction and energy becomes a daunting task even for the smallest system. An alternative route to the direct solution of the Schrödinger equation is shifting focus of the problem from  $\Psi(\vec{r})$  to a fundamental observable of the problem, namely the density  $n(\vec{r})$ . For a many-particle system the density,  $n(\vec{r})$ , is found by calculating the expectation value of the single-particle density operator for the many-body wavefunction

$$\hat{n}(\vec{r}) = \sum_{i=1,N} \delta(\vec{r} - \vec{r}_i) \quad (1)$$

$$\begin{aligned} n(\vec{r}) &= \langle \Psi | \hat{n}(\vec{r}) | \Psi \rangle = \sum_{i=1,N} \int \delta(\vec{r} - \vec{r}_i) |\Psi(\vec{r}_1, \dots, \vec{r}_N)|^2 d\vec{r}_1 \dots d\vec{r}_N \\ &= \int |\Psi(\vec{r}, \vec{r}_2, \dots, \vec{r}_N)|^2 d\vec{r}_2 d\vec{r}_3 \dots d\vec{r}_N + \int |\Psi(\vec{r}_1, \vec{r}, \dots, \vec{r}_N)|^2 d\vec{r}_1 d\vec{r}_3 \dots d\vec{r}_N + \dots \\ &= N \int |\Psi(\vec{r}, \dots, \vec{r}_N)|^2 d\vec{r}_2 \dots d\vec{r}_N \end{aligned} \quad (2)$$

where  $\{\vec{r}_i\}$  are the variable associated with each of the electrons. As expected, assuming that the wavefunction is normalized to unity, integrating the electron density over all space yields the number of electrons

$$\int d\vec{r} n(\vec{r}) = N. \quad (3)$$

## II. ENERGY IN TERMS OF THE DENSITY

Our goal in density functional theory is to eliminate the wavefunction by writing all terms making up the total ground state energy of the electronic system in terms of density. This is because we would then like to minimize the electronic energy with respect to the density to obtain the ground state energy and the corresponding electronic density. The foundation put forward by Hohenberg and Kohn will be discussed in the next lecture.

As is well-known, once the wavefunction is obtained by solving the Hamiltonian we can determine the observable corresponding to a given operator by calculating the expectation value of that operator. This allows us to separately calculate the energy terms corresponding to the potential operators given in the Hamiltonian that we wrote down in the previous lectures. For the sake of completeness we reproduce the Hamiltonian here once again

$$\begin{aligned} \hat{H}_e &= \hat{T} + \hat{V}_{en} + \hat{V}_{ee} \\ &= -\frac{1}{2} \sum_i^{N_e} \nabla_i^2 - \sum_i^{N_e} \sum_I^{N_n} \frac{Z_I}{|\vec{r}_i - \vec{R}_I|} + \frac{1}{2} \sum_i^{N_e} \sum_{j \neq i}^{N_e} \frac{1}{|\vec{r}_i - \vec{r}_j|} \end{aligned} \quad (4)$$

Suppose now that we have somehow managed to solve the many-body Schrödinger equation and have obtained the wavefunction. The expectation value of the nuclei-electron interaction operator is given by

$$\langle \Psi(\vec{r}_1, \dots, \vec{r}_N) | \hat{V}_{ne} | \Psi(\vec{r}_1, \dots, \vec{r}_N) \rangle = - \sum_i^{N_e} \sum_I^{N_n} \int \Psi^*(\vec{r}_1, \dots, \vec{r}_N) \frac{Z_I}{|\vec{r}_i - \vec{R}_I|} \Psi(\vec{r}_1, \dots, \vec{r}_N) d\vec{r}_1 \dots d\vec{r}_N \quad (5)$$

Since the operator does not contain any derivatives, we may collect the wavefunction and its conjugate under a common norm square

$$\langle \Psi(\vec{r}_1, \dots, \vec{r}_N) | \hat{V}_{ne} | \Psi(\vec{r}_1, \dots, \vec{r}_N) \rangle = - \sum_i^{N_e} \sum_I^{N_n} \int \frac{Z_I}{|\vec{r}_i - \vec{R}_I|} |\Psi(\vec{r}_1, \dots, \vec{r}_N)|^2 d\vec{r}_1 \dots d\vec{r}_N \quad (6)$$

Next, we proceed in a way that is much similar to the one we followed in the calculation of the density. Let's expand the sum over the electronic index  $i$

$$\langle \Psi(\vec{r}_1, \dots, \vec{r}_n) | \hat{V}_{ne} | \Psi(\vec{r}_1, \dots, \vec{r}_n) \rangle = - \sum_I^{N_n} \left[ \int \frac{Z_I}{|\vec{r}_1 - \vec{R}_I|} |\Psi(\vec{r}_1, \dots, \vec{r}_N)|^2 d\vec{r}_1 \dots d\vec{r}_N \right. \\ \left. + \int \frac{Z_I}{|\vec{r}_2 - \vec{R}_I|} |\Psi(\vec{r}_1, \dots, \vec{r}_N)|^2 d\vec{r}_1 \dots d\vec{r}_N + \dots \right] \quad (7)$$

For each of the  $N_e$  terms in the sum in Eq. 7, we separate the integral over the variable in the Coulombic terms from the others.

$$\langle \Psi(\vec{r}_1, \dots, \vec{r}_n) | \hat{V}_{ne} | \Psi(\vec{r}_1, \dots, \vec{r}_n) \rangle = - \sum_I^{N_n} \left[ \int \frac{Z_I}{|\vec{r}_1 - \vec{R}_I|} d\vec{r}_1 \int |\Psi(\vec{r}_1, \dots, \vec{r}_N)|^2 d\vec{r}_2 d\vec{r}_3 \dots d\vec{r}_N \right. \\ \left. + \int \frac{Z_I}{|\vec{r}_2 - \vec{R}_I|} d\vec{r}_2 \int |\Psi(\vec{r}_1, \dots, \vec{r}_N)|^2 d\vec{r}_1 d\vec{r}_3 \dots d\vec{r}_N + \dots \right] \quad (8)$$

For each term in Eq. 8, the second integral is the definition of the density,  $n(\vec{r})$  as given in Eq. 2, that is to say Eq. 8 may be written as

$$E_{ne} = \langle \Psi(\vec{r}_1, \dots, \vec{r}_n) | \hat{V}_{ne} | \Psi(\vec{r}_1, \dots, \vec{r}_n) \rangle = - \frac{1}{N_e} \sum_I^{N_n} \left[ \int \frac{Z_I}{|\vec{r}_1 - \vec{R}_I|} n(\vec{r}_1) + \int \frac{Z_I}{|\vec{r}_2 - \vec{R}_I|} n(\vec{r}_2) + \dots \right] \quad (9)$$

Since for each term in Eq. 9, an integral over the argument of the density is taken, we may replace these variables by a dummy variable, say  $vr$ . Then all the terms become equal and since there are  $N_e$  of them the electron-nucleus interaction energy may be written takes the compact form

$$E_{ne} = - \sum_I^{N_n} \int n(\vec{r}) \frac{Z_I}{|\vec{r} - \vec{R}_I|} d\vec{r} = \int n(\vec{r}) V_{ne}(\vec{r}) d\vec{r} \quad (10)$$

This simple linear functional form attained by the electron-nucleus interaction energy is in fact common to all single-body external potential terms such as an applied electric field.

The equivalent derivation for the electron-electron term, however, is not as simple. It can be proven that this term cannot be written in terms of the single-particle density but instead only in terms of the two-particle density

$$E_{ee} = \frac{1}{2} \int \int d\vec{r} d\vec{r}' \frac{n^{(2)}(\vec{r}, \vec{r}')}{|\vec{r} - \vec{r}'|} \quad (11)$$

where  $n^{(2)}$  may be interpreted as the probability that an electron exists at point  $\vec{r}$  given that a second electron exists at point  $\vec{r}'$ . Eq. 11 may easily be derived in a similar manner to  $E_{ne}$ . It is this term that makes the many-particle problem so hard to solve. If we knew this pairwise conditional probability, then we would be able to solve the problem exactly. However, the method that we are about to develop does not allow for the existence of such a two-particle density but instead only the one-particle density. That being the case, we are forced to make an approximation. Now, if the two electrons were completely *uncorrelated* then the two-particle density in Eq. 11 would just be the product of one-particle densities (we may also make the same statement about probabilities). Here, we write  $n^{(2)}$  in terms of this product plus a correction

$$n^{(2)}(\vec{r}, \vec{r}') = n(\vec{r})n(\vec{r}') + \Delta n^{(2)}(\vec{r}, \vec{r}') \quad (12)$$

The electron-electron energy of Eq. 11 may then be written as

$$E_{ee} = \frac{1}{2} \int \int d\vec{r} d\vec{r}' \frac{n(\vec{r})n(\vec{r}')}{|\vec{r} - \vec{r}'|} + \Delta E_{ee} \quad (13)$$

where the second term obviously comes from the correction in Eq. 12.

The kinetic energy poses an even larger problem. Since the kinetic energy operator contains a derivative term, there's no way we can write it in terms of the density since it is impossible to collect the wavefunction and its conjugate as a single norm square.

$$T = - \frac{1}{2} \int d\vec{r} \Psi^*(\vec{r}_1, \dots, \vec{r}_n) \nabla^2 \Psi(\vec{r}_1, \dots, \vec{r}_n) \quad (14)$$

In order to tackle the kinetic energy, we make one of the key assumptions of density functional theory. We assume that the density can be written as the sum norm squares of a collection of single-particle orbitals

$$n(\vec{r}) = \sum_n^{N_e} |\phi_n(\vec{r})|^2 \quad (15)$$

These orbitals are called *Kohn-Sham* orbitals and they are initially completely unspecified in much the same way as in the orbitals in the Slater determinant in the Hartree-Fock formalism. The above form cannot really be considered an approximation. It simply says that instead of the full many-particle system we consider an auxiliary system of single-particle orbitals that have the same ground state density as the real system. Whether such an auxiliary system can be found for every real system is a topic of debate but if it can be found the treatment remains exact.

Writing the density as we have in Eq. 15 suggests that we write the kinetic energy as the sum of the kinetic energies of the Kohn-Sham orbitals. However, this of course is not going to equal the kinetic energy of the real many-particle system. Once again then we express the kinetic energy as the single-particle kinetic energy plus a correction

$$T = -\frac{1}{2} \sum_n^{N_e} \int d\vec{r} \phi_n^*(\vec{r}) \nabla^2 \phi_n(\vec{r}) d\vec{r} + \Delta T \quad (16)$$

Finally putting together everything, the total ground state energy may be written as

$$E = -\frac{1}{2} \sum_n^{N_e} \int d\vec{r} \phi_n^*(\vec{r}) \nabla^2 \phi_n(\vec{r}) d\vec{r} + \int n(\vec{r}) V_{ne}(\vec{r}) d\vec{r} + \frac{1}{2} \int \int d\vec{r} d\vec{r}' \frac{n(\vec{r}) n(\vec{r}')}{|\vec{r} - \vec{r}'|} + \Delta T + \Delta E_{ee} \quad (17)$$

The kinetic energy as written cannot be differentiated directly with respect to  $n(\vec{r})$  but it can be minimized with respect to the orbitals which is equivalent to minimizing with respect to  $n(\vec{r})$ .

Now the only thing that remains to be done is to write the last two (correction) terms in terms of the density as well. In density functional theory, these additional terms,  $\delta E_{ee}$  and  $\Delta E_K$  play a crucial role. So far, our conversion of the energy from the wavefunction representation to the density representation has been exact. From this point forward, we will rely on the accuracy of certain approximations to the sum of these two terms, called the *exchange-correlation* energy.

$$E_{xc} = \Delta E_{ee} + \Delta T \quad (18)$$

The origin of this term is the difference between a system of  $N$  interacting and noninteracting particles. More specifically, the origin of

- exchange energy is the Pauli repulsion, omitted in the Hartree term.
- correlation energy is the repulsion between electrons.

Several well-developed approximations to this sum exists but here we shall only consider the local approximation given by

$$E_{xc} = \int d\vec{r} n(\vec{r}) \epsilon_{xc}(n) \quad (19)$$

where  $\epsilon_{xc}(n)$  is a simple function of  $n$ . We will see more on exchange-correlation functionals later. Thus, within this approximation, the total energy may be written as

$$E = -\frac{1}{2} \sum_n^{N_e} \int d\vec{r} \phi_n^*(\vec{r}) \nabla^2 \phi_n(\vec{r}) d\vec{r} + \int n(\vec{r}) V_{ne}(\vec{r}) d\vec{r} + \frac{1}{2} \int \int d\vec{r} d\vec{r}' \frac{n(\vec{r}) n(\vec{r}')}{|\vec{r} - \vec{r}'|} + \int d\vec{r} \epsilon_{xc}(n) n(\vec{r}). \quad (20)$$

We'll use this result first to prove the most important theorems (by Hohenberg and Kohn) on DFT and later to use functional derivatives to obtain the Kohn-Sham system of equations which make DFT possible in practice.

## Lecture VII : The Hohenberg-Kohn theorem and the Kohn-Sham equations

### I. THE HOHENBERG-KOHN THEOREM

DFT is made possible by the existence of two ingeniously simple theorems put forward and proven by Hohenberg and Kohn in 1964 (see *Physical Review*, **136**, B864) .

**Theorem I :** For any system of interacting particles in an external potential  $V_{ext}(\vec{r})$ , the density is uniquely determined (in other words, the external potential is a unique *functional* of the density).

**Proof I :** Assume that there exist two potentials  $V_{ext}^{(1)}(\vec{r})$  and  $V_{ext}^{(2)}(\vec{r})$  differing by more than a constant and giving rise to the same ground state density,  $n(\vec{r})$ . Obviously,  $V_{ext}^{(1)}(\vec{r})$  and  $V_{ext}^{(2)}(\vec{r})$  belong to distinct Hamiltonians  $\hat{H}_{ext}^{(1)}(\vec{r})$  and  $\hat{H}_{ext}^{(2)}(\vec{r})$ , which give rise to distinct wavefunctions  $\Psi_{ext}^{(1)}(\vec{r})$  and  $\Psi_{ext}^{(2)}(\vec{r})$ .

Because of the variational principle, no wavefunction can give an energy that is less than the energy of  $\Psi_{ext}^{(1)}(\vec{r})$  for  $\hat{H}_{ext}^{(1)}(\vec{r})$ . That is

$$E^{(1)} = \langle \Psi^{(1)} | \hat{H}^{(1)} | \Psi^{(1)} \rangle < \langle \Psi^{(2)} | \hat{H}^{(1)} | \Psi^{(2)} \rangle. \quad (1)$$

Assuming that the ground state is not degenerate, the inequality strictly holds. Because we have identical ground state densities for the two Hamiltonians, we can rewrite the expectation value in Eq. 1 as

$$\langle \Psi^{(2)} | \hat{H}^{(1)} | \Psi^{(2)} \rangle = \langle \Psi^{(2)} | \hat{H}^{(2)} | \Psi^{(2)} \rangle + \int d\vec{r} [V_{ext}^{(1)}(\vec{r}) - V_{ext}^{(2)}(\vec{r})] n_0(\vec{r}) \quad (2)$$

Exchanging labels in Eq. 2, we obtain

$$\langle \Psi^{(1)} | \hat{H}^{(2)} | \Psi^{(1)} \rangle = \langle \Psi^{(1)} | \hat{H}^{(1)} | \Psi^{(1)} \rangle + \int d\vec{r} [V_{ext}^{(2)}(\vec{r}) - V_{ext}^{(1)}(\vec{r})] n_0(\vec{r}) \quad (3)$$

Adding Eq. 2 and Eq. 3, we obtain

$$E^{(1)} + E^{(2)} < E^{(2)} + E^{(1)} \quad (4)$$

which is clearly a contradiction. Thus, the theorem has been proven by *reductio ad absurdum*.

**Theorem II :** A universal *functional* for the energy  $E[n]$  can be defined in terms of the density. The exact ground state is the global minimum value of this functional.

**Proof II :** Since the external potential is uniquely determined by the density and since the potential in turn uniquely (except in degenerate situations) determines the ground state wavefunction, all the other observables of the system such as kinetic energy are uniquely determined. Then one may write the energy as a functional of the density. We have in fact already shown in detail in the previous lecture how to convert the total energy expression from wavefunction representation to density representation.

$$E[n] = T[n] + E_{int}[n] + \int V_{ext}(\vec{r})n(\vec{r}) + E_{II} \equiv F[n] + \int V_{ext}(\vec{r})n(\vec{r}) + E_{II} \quad (5)$$

where  $F[n]$  is a universal functional because the treatment of the kinetic and internal potential energies are the same for all systems.

In the ground state the energy is defined by the unique ground state density,  $n^{(1)}(\vec{r})$ ,

$$E^{(1)} = E[n^{(1)}] = \langle \Psi^{(1)} | \hat{H}^{(1)} | \Psi^{(1)} \rangle. \quad (6)$$

By the variational principle, a different density,  $n^{(2)}(\vec{r})$  will necessarily give a higher energy

$$E^{(1)} = E[n^{(1)}] = \langle \Psi^{(1)} | \hat{H}^{(1)} | \Psi^{(1)} \rangle < \langle \Psi^{(2)} | \hat{H}^{(1)} | \Psi^{(2)} \rangle = E^{(2)}. \quad (7)$$

It follows that minimizing with respect to  $n(\vec{r})$  the total energy of the system written as a functional of  $n(\vec{r})$ , one finds the total energy of the ground state. The correct density that minimizes the energy is then the ground state density.



## II. SOME IMPORTANT ISSUES RELATED TO HOHENBERG-KOHN THEOREMS

As simple as they sound, Hohenberg and Kohn theorems raised many questions regarding the extraordinary claim that there is a one-to-one matching between densities and external potentials. In fact, it turns out that the Hohenberg and Kohn theorems apply under certain circumstances. Fortunately, we usually do not concern ourselves with these points since in practice, densities of atomic systems do obey these constraints. However, for the sake of completeness we shall now consider some of these fundamental issues.

### A. Degenerate ground states

If there are more than one wavefunctions corresponding to the ground state energy of a system, one can no longer talk about the uniqueness of the ground state expectation value of operators. In particular, one cannot prove the first theorem. In the next subsection, we will see, in connection with another problem, a different way of formulating the above problem, which will also resolve the degeneracy issue as a side effect.

### B. N- and V-representability, Lieb-Levy constrained search formalism

A *V-representable* density is a ground state density that can be associated with a Hamiltonian that has an external potential,  $v(\vec{r})$ . In the Hohenberg-Kohn theorems, one important assumption which is not mentioned is that, during the minimization, as we vary the density, we assume that it remains V-representable. It is, however, not clear that an arbitrary density, which integrates to an integer number of electrons, would be the ground state of a smooth external potential.

**Example :** Any excited state density of single particles in finite systems.

In order to overcome the V-representability problem, a more general variational routine was proposed independently by Levy and Lieb in their respective pioneering works. Their extended minimization algorithm required the densities to be only *N-representable*. An N-representable density is one which can be derived from an antisymmetric wavefunction. This condition, being much weaker than V-representability, was in fact proven by Harriman to be easily satisfied by an arbitrary density.

**Motivation for the proof of Harriman's treatment :** For a one-dimensional density  $n(x)$  that integrates to  $N$ , define the auxiliary function  $f(x)$  as

$$\frac{df(x)}{dx} = \frac{2\pi}{N}n(x) \quad (8)$$

which through integration yields

$$f(x) = \frac{2\pi}{N} \int_{-\infty}^x n(x')dx'. \quad (9)$$

Next define a set of single particle orbitals

$$\phi_k(x) = \sqrt{\frac{n(x)}{N}} e^{i[kf(x) + \phi(x)]} \quad (10)$$

where  $k$  is an integer and  $\phi(x)$  is an arbitrary phase factor. The orbitals such defined can easily be demonstrated to be orthogonal and complete (your next homework). A Slater orbital,  $\Psi_{k_1, \dots, k_N}$  made up of  $N$  such orbitals may always be constructed. Using the fact that the density operator is a single-particle operator, we may immediately use our knowledge from Hartree-Fock derivations to obtain

$$\langle \Psi_{k_1, \dots, k_N} | \hat{n} | \Psi_{k_1, \dots, k_N} \rangle = \sum_{i=1}^N |\phi_{k_i}|^2 = n(x) \quad (11)$$

where the last equation has been obtained using Eq. 10.

Following the 1979 paper by Levy (PNAS, **(76)**, 6062), we first redefine the *universal* part of the Hohenberg-Kohn energy functional. Given a Hamiltonian

$$\hat{H} = \hat{T} + \hat{V}_{ee} + \sum_i \hat{v}_i \quad (12)$$

$F[n]$  was defined by Hohenberg and Kohn to be the sum of the kinetic and the electron-electron interaction energies. In the new definition of Levy and Lieb, the universal part of the energy functional

$$Q[n] = \min_{\Psi_n \rightarrow n} \langle \Psi_n | \hat{T} + \hat{V}_{ee} | \Psi_n \rangle \quad (13)$$

where a *constrained minimization* is conducted over all the antisymmetric wavefunctions  $\Psi_n$ , which yield the trial density  $n$ . With the definition in Eq. 13, of course, we assume that the arbitrary trial density can be obtained from an antisymmetric wavefunction, in other words, is N-representable.

In order for  $Q[n]$  to be a valid universal function, it must satisfy similar conditions as set forth by the original Hohenberg-Kohn theorems. In contrast to the Hohenberg-Kohn theorems, Lieb-Levy theorems are not almost trivially proven.

**Theorem I :** For an arbitrary N-representable density  $n$ ,

$$\int d\vec{r} \hat{v}(\vec{r}) n(\vec{r}) + Q[n] \geq E_0 \quad (14)$$

where  $\hat{v}$  is the one-body operator which constitutes the external potential

$$\hat{V}_{ext} = \sum_i \hat{v}(\vec{r}_i) \quad (15)$$

and  $E_0$  is the ground state energy.

**Proof :** Following the notation of Levy in his PNAS paper, define  $\Psi_{min}^n$  as the wavefunction that minimizes  $\langle \Psi^n | \hat{T} + \hat{V}_{ee} | \Psi^n \rangle$  for a fixed  $n$ . Then by the definition in Eq. 13, we have

$$Q[n] = \langle \Psi_{min}^n | \hat{T} + \hat{V}_{ee} | \Psi_{min}^n \rangle \quad (16)$$

and in particular for the ground state density that we seek

$$Q[n_{gs}] = \langle \Psi_{min}^{n_{gs}} | \hat{T} + \hat{V}_{ee} | \Psi_{min}^{n_{gs}} \rangle. \quad (17)$$

The minimum total energy for an arbitrary density,  $n$  (the left-hand side of the inequality in Eq. 14) is then given by

$$\int d\vec{r} \hat{v}(\vec{r}) n(\vec{r}) + Q[n] = \int d\vec{r} \hat{v}(\vec{r}) n(\vec{r}) + \langle \Psi_{min}^n | \hat{T} + \hat{V}_{ee} | \Psi_{min}^n \rangle \quad (18)$$

which can alternatively be expressed as

$$\int d\vec{r} \hat{v}(\vec{r}) n(\vec{r}) + Q[n] = \langle \Psi_{min}^n | \hat{T} + \hat{V}_{ee} + \hat{V}_{ext} | \Psi_{min}^n \rangle \quad (19)$$

making use of Eq. 15.

Since, in the general case, the density  $n$  in Eq. 19 is not the ground state density we have, by the variational principle,

$$\langle \Psi_{min}^n | \hat{T} + \hat{V}_{ee} + \hat{V}_{ext} | \Psi_{min}^n \rangle \geq E_0. \quad (20)$$

Combining Eq. 19 and Eq. 20, we have

$$\int d\vec{r} \hat{v}(\vec{r}) n(\vec{r}) + Q[n] \geq E_0 \quad (21)$$

which completes the proof of the first theorem.

**Theorem II :** For the ground state density,

$$\int d\vec{r} \hat{v}(\vec{r}) n_{gs}(\vec{r}) + Q[n_{gs}] = E_0 \quad (22)$$

**Proof :** Relying on the considerations illustrated so far, the true ground state density of the system  $\Psi_{gs}$  is not necessarily equal to the wavefunction that minimizes  $Q[n_{gs}]$ , i.e.  $\Psi_{min}^{n_{gs}}$ . As a result, the variational principle dictates that

$$\langle \Psi_{gs} | \hat{T} + \hat{V}_{ee} + \hat{V}_{ext} | \Psi_{gs} \rangle \leq \langle \Psi_{min}^{n_{gs}} | \hat{T} + \hat{V}_{ee} + \hat{V}_{ext} | \Psi_{min}^{n_{gs}} \rangle. \quad (23)$$

Eq. 23 can be rewritten by separating the external potential energy term

$$\int d\vec{r} |\Psi_{gs}|^2 \hat{v}(\vec{r}) + \langle \Psi_{gs} | \hat{T} + \hat{V}_{ee} | \Psi_{gs} \rangle \leq \int d\vec{r} n_{gs}(\vec{r}) \hat{v}(\vec{r}) + \langle \Psi_{min}^{n_{gs}} | \hat{T} + \hat{V}_{ee} | \Psi_{min}^{n_{gs}} \rangle \quad (24)$$

which, making use of the fact that the first terms on both sides gives the same result, can be reduced to

$$\langle \Psi_{gs} | \hat{T} + \hat{V}_{ee} | \Psi_{gs} \rangle \leq \langle \Psi_{min}^{n_{gs}} | \hat{T} + \hat{V}_{ee} | \Psi_{min}^{n_{gs}} \rangle. \quad (25)$$

On the other hand, since both wavefunctions correspond to the same density, namely the ground state density, variational principle also dictates

$$\langle \Psi_{gs} | \hat{T} + \hat{V}_{ee} | \Psi_{gs} \rangle \geq \langle \Psi_{min}^{n_{gs}} | \hat{T} + \hat{V}_{ee} | \Psi_{min}^{n_{gs}} \rangle. \quad (26)$$

The only way Eq. 25 and Eq. 26 can both be satisfied is if equality holds. In other words,

$$\langle \Psi_{gs} | \hat{T} + \hat{V}_{ee} | \Psi_{gs} \rangle = Q[n_{gs}]. \quad (27)$$

Combined with the definition of the ground state energy using the true ground state wavefunction

$$E_0 = \int d\vec{r} \hat{v}(\vec{r}) n_{gs}(\vec{r}) + \langle \Psi_{gs} | \hat{T} + \hat{V}_{ee} | \Psi_{gs} \rangle \quad (28)$$

Eq. 27 yields

$$\int d\vec{r} \hat{v}(\vec{r}) n_{gs}(\vec{r}) + Q[n_{gs}] = E_0 \quad (29)$$

and completes the proof of the second theorem.

We have thus replaced the V-representability condition required by the Hohenberg-Kohn theorems by the much weaker N-representability condition. In addition to this commodity, the Lieb-Levy constrained search has two very desirable side effects :

- As put forward by the second theorem,  $F[n]$  in Eq. 12 is equal to Eq. 13 when the density is V-representable, or in other words the Lieb-Levy search is equivalent to the Hohenberg-Kohn minimization. This, combined with the fact that non V-representable densities are not commonly found in physical problems enables us to just go ahead and use the regular, Hohenberg-Kohn formalism without worrying about the fundamental applicability.
- Since  $\Psi_{gs} = \Psi_{min}^{gs}$ , the ground state wavefunction may be found in terms of the ground state density even if the external potential is not known. (This is in fact another statement of N-representability.) One would need only to go over the wavefunctions yielding the ground state density and find the one that gives minimum expectation value of  $\hat{T} + \hat{V}_{ee}$ . As such, there is no restriction on the number of wavefunctions that would satisfy this condition. Therefore, this freedom also resolves the important issue of degeneracy, identifying (in principle) all the degenerate ground state densities.

### C. Fractional Total Particle Number, Discontinuous Derivatives

The densities considered throughout the minimization of the energy functional must strictly obey the particle number condition

$$\int d\vec{r} n(\vec{r}) = N. \quad (30)$$

As is the usual practice, such a constraint can be introduced into the minimization scheme through Lagrange multipliers. Then the minimization proceeds as

$$\frac{\delta [E_v[n] - \mu \int d\vec{r} n(\vec{r})]}{\delta n} = 0 \quad (31)$$

where the subscript  $v$  denotes an energy that is related to the external potential constructed through  $\hat{v}$  and  $\mu$  is the Lagrange multiplier. With the Lagrange multiplier taking into account the constraint explicitly, we extend the minimization to an arbitrary number of particles. In fact, this number is even allowed to be fractional. Of course, a fractional number of

electrons have no further significance than simply being a mathematical tool. In the case of fractional total particle number, Eq. 30 can be rewritten as (with the notation taken from Dreizler and Gross, *density Functional Theory : An Approach to the Quantum Many-Body Problem*)

$$\int d\vec{r} n(\vec{r}) = M + \omega \quad (32)$$

where  $M$  is an integer and  $\omega$  is a real number between 0 and 1. A suitable definition of a universal functional expressed in terms of a *mixture* of an  $M$ -particle density and an  $(M + 1)$ -particle density

$$F[n] = \min_{n_{\text{mix}}} \langle \Psi_{\text{mix}} | \hat{T} + \hat{V}_{ee} | \Psi_{\text{mix}} \rangle \quad (33)$$

where the minimization goes over all possible mixtures

$$n_{\text{mix}} = \alpha_M \langle \Psi_M | \hat{n}(\vec{r}) | \Psi_M \rangle + \alpha_{M+1} \langle \Psi_{M+1} | \hat{n}(\vec{r}) | \Psi_{M+1} \rangle. \quad (34)$$

Integrating  $n_{\text{mix}}$  over space and comparing to Eq. 32, we arrive at the rather obvious result

$$\alpha_{M+1} = \omega \quad \text{and} \quad \alpha_M = (1 - \omega). \quad (35)$$

With these definitions, we are allowed to carry out the minimization with the Lagrange multiplier yielding

$$\frac{\partial E_v[n]}{\partial n} = \mu. \quad (36)$$

As is mostly the case when using Lagrange multipliers, we now try to see if  $\mu$  has a physical meaning. In particular, we look at the derivative of the total energy with respect to the number of particles, which is known to be the chemical potential.

$$\frac{\partial E_N}{\partial N} = \lim_{\epsilon \rightarrow 0} \frac{E_{N+\epsilon} - E_N}{\epsilon} = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \int d\vec{r} \frac{\delta E_v[n]}{\delta n(\vec{r})} \Big|_{n_N} (n_{N+\epsilon} - n_N) = \frac{1}{\epsilon} \mu \int (n_{N+\epsilon} - n_N) d\vec{r} = \frac{1}{\epsilon} \mu \epsilon = \mu \quad (37)$$

where the last equations have been obtained by substituting Eq. 36 into Eq. 37. We see then that our Lagrange multiplier is indeed the chemical potential.

Next, we question the nature of the chemical potential  $\mu(N)$  as a function of  $N$  and reach an interesting conclusion that has serious implications on the prediction of solid state properties.

Assume  $\mu(N)$  is a continuous function of  $N$ . If we have two atoms that have nuclear charges  $X$  and  $Y$  and if they are infinitely separated, then the total energy is a sum of the individual energies of the two isolated atoms

$$E_{\text{tot}} = E(X) + E(Y). \quad (38)$$

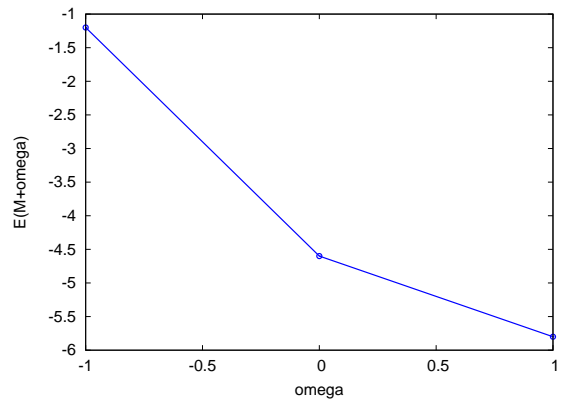
Now assume that  $\mu(X) < \mu(Y)$ . If we now transfer an infinitesimal amount of charge,  $\delta N_x$  from  $Y$  to  $X$ , the change in energy is going to be  $(\mu(X) - \mu(Y))\delta N_x$ , which is negative. Thus, we end up with the unreasonable situation that the separated ionic configuration is favorable over the separated neutral atoms. This result however is a direct consequence of the assumption of a continuous chemical potential. On the contrary, going back to the fractional minimization scheme of Eq. 33, the minimum energy for a fractional number of particles is (in the Lieb-Levy formalism)

$$E_{M+\omega} = \min_{n(\vec{r}) \rightarrow N+\omega} \min_{\Psi_M, \Psi_{M+1} \rightarrow n(\vec{r})} \left[ (1 - \omega) \langle \Psi_M | \hat{H} | \Psi_M \rangle + \omega \langle \Psi_{M+1} | \hat{H} | \Psi_{M+1} \rangle \right] \quad (39)$$

where the inner minimization goes over all combinations of  $M$  and  $(M + 1)$ -particle asymmetric wavefunctions that yield a given density  $n(\vec{r})$  and the outer one points to the minimum of all such densities that integrate to  $N + \omega$  particles. The smallest possible value of this energy is found when both the  $M$  and the  $M + 1$ -particle systems attain their ground state, namely

$$E_{M+\omega} = (1 - \omega)E_M + \omega E_{M+1}. \quad (40)$$

Considering now three consecutive energies  $E_{M-1}$ ,  $E_M$  and  $E_{M+1}$ ,  $E_{M+\omega}$  would look similar to the graph to



the right, with discontinuities at integral number of particles. This also implies that the same discontinuities will occur in the functional derivative of the energy in Eq. 36.

This statement has important implications which will be detailed later. This discontinuity also resolves the separated atoms paradox illustrated above. The ionization potential and the electron affinity are defined as

$$I(Z) = E_{Z-1}(Z) - E_Z(Z) \quad (41)$$

$$A(Z) = E_Z(Z) - E_Z(Z+1) \quad (42)$$

The cost of transferring one electron from  $Y$  to  $X$  then will be given by

$$\begin{aligned} \Delta E_{Y \rightarrow X} &= \underbrace{E_{X+1}(X) + E_Y(Y-1)}_{\text{final state}} - \underbrace{E_X(X) + E_Y(Y)}_{\text{initial state}} \\ &= E_{X+1}(X) - E_X(X) + E_{Y-1}(Y) - E_Y(Y) = -A(X) + I(Y). \end{aligned} \quad (43)$$

Conversely, the cost of moving an electron from  $X$  to  $Y$  will be given by

$$\Delta E_{X \rightarrow Y} = -A(Y) + I(X). \quad (44)$$

Since the largest electron affinity in nature (Cl) is smaller than the smallest ionization potential (Cs), both  $\Delta E_{Y \rightarrow X}$  and  $\Delta E_{X \rightarrow Y}$  must be greater than zero. That is to say any situation that is different from two neutral separated atoms must result in a larger energy. This can only be possible if the chemical potential has discontinuities at integral values of total energy, with a minimum at zero charge transfer.

### III. KOHN-SHAM EQUATIONS

Utilizing the Hohenberg-Kohn theorems, we minimize the total energy with respect to the orbitals in order to obtain the orbitals that give rise to the ground state energy. While performing the minimization, we prefer to minimize with respect to  $\phi_i^*(\vec{r})$  instead of  $\phi_i(\vec{r})$ . One can prove that both yield the same result.

Just like regular differentiation, we can employ chain rule for functional derivatives. This of course works for all the terms except for kinetic energy. Kinetic energy may be differentiated directly with respect to the orbital. We thus have

$$\frac{\delta E_e}{\delta \phi_i^*(\vec{r})} = \frac{\delta T_s}{\delta \phi_i^*(\vec{r})} + \left[ \frac{\delta E_{ext}}{\delta n(\vec{r})} + \frac{\delta E_{Hartree}}{\delta n(\vec{r})} + \frac{\delta E_{xc}}{\delta n(\vec{r})} \right] \frac{\delta n(\vec{r})}{\delta \phi_i^*(\vec{r})} = \varepsilon_i \phi_i(\vec{r}) \quad (45)$$

$$-\frac{1}{2} \nabla^2 \phi_i(\vec{r}) + \left[ V_{ext}(\vec{r}) + \underbrace{\int d\vec{r}' \frac{n(\vec{r}')}{|\vec{r} - \vec{r}'|}}_{V_H} + \underbrace{\epsilon_{xc}[n] + n(\vec{r}) \frac{\delta \epsilon_{xc}[n]}{\delta n(\vec{r})}}_{V_{xc}} \right] \phi_i(\vec{r}) = \varepsilon_i \phi_i(\vec{r}) \quad (46)$$

Eq. 46 is in fact a system of equations, which when solved simultaneously represent the many-particle system in terms of single-particle orbitals. Each of these equations resemble a Schrödinger equation

$$[\hat{T} + V_{eff}] \phi_i(\vec{r}) = \varepsilon_i \phi_i(\vec{r}), \quad (47)$$

with the important difference that  $V_{eff}$  which we have defined to be the sum of the terms  $V_H$ ,  $V_{xc}$  and  $V_{ext}$ , depends on the density and indirectly on the orbitals. As a result we have the unusual situation that any change in the orbitals effect also the potential on which they in turn depend. This chicken-or-egg character is resolved by solving the Kohn-Sham system of equations self-consistently.

### IV. ISSUES REGARDING THE KOHN-SHAM METHODS

#### A. Meaning of Kohn-Sham eigenvalues

Once again, as in the Hartree-Fock treatment, we end up with a many-body system that is described by a set of single-particle orbitals. The sum of the Kohn-Sham eigenvalues can be obtained by multiplying Eq. 46 from the left by  $\phi_i^*(\vec{r})$ ,

integrating over space and summing over  $i$ . This sequence of operations yield

$$\sum \epsilon_i = T_s + \int d\vec{r} n(\vec{r}) V_{ext} + \int \int d\vec{r} d\vec{r}' \frac{n(\vec{r})n(\vec{r}')}{|\vec{r} - \vec{r}'|} + \int d\vec{r} n(\vec{r}) \epsilon_{xc}(n(\vec{r})) + \int d\vec{r} n^2(\vec{r}) \frac{\partial \epsilon_{xc}}{\partial n} \quad (48)$$

which differs from the total energy functional derived in the previous lecture by

$$\Delta E = \sum_i \epsilon_i - E_{tot}[n] = \frac{1}{2} E_H + \int d\vec{r} n^2(\vec{r}) \frac{\partial \epsilon_{xc}}{\partial n}. \quad (49)$$

The eigenvalues thus do not have a significance as single particle energies as expected since the description of the system as a collection of single particles is incorrect.

Curiously enough, however, the Kohn-Sham orbitals have proven to give quite accurate descriptions of band structures and bonding characters. There has been an ongoing investigation of why this happens and whether Kohn-Sham orbitals carry any significance (See for instance Stowasser and Hoffmann, JACS, 1999, **121**, 3414).

## B. Fractional Occupations and Janak's Theorem

The meaning of the Kohn-Sham orbitals is itself a very prominent question in DFT theory. While in theory they should not have any physical meaning, empirically it has been observed over and over again that the eigenvalues give a reasonable description of the band structure. The eigenvalues in the Hartree-Fock theory was found to represent the energy necessary to add or remove an electron from a given orbital (assuming the correctness of the single-particle description) by Koopmans' theorem. A similar theorem was developed by Janak in 1978 (Phys Rev B **18**, 7165). The title of the paper "*Proof that  $\frac{\partial E}{\partial n_i} = \epsilon_i$  in density-functional theory*" is also the main result of his treatment.

In order to search for the meaning (at least mathematically) of the Kohn-Sham eigenvalues,  $\epsilon_i$ , the necessity of defining fractional total particle number in the execution of Hohenberg-Kohn theorems resurfaces in the form of fractional occupations. To this end, we first give an equivalent description of the density to the conventional one, incorporating also the occupations

$$n(\vec{r}) = \sum_{i=1}^N |\psi_i|^2 = \sum_{i=1}^{\infty} n_i |\psi_i|^2 \quad (50)$$

where following Janak's original notation  $n_i$  are the occupations of the Kohn-Sham orbitals. Of course the ground state density is the one that corresponds to the occupation of the  $N$  lowest Kohn-Sham orbitals, i.e. the Fermi-Dirac distribution at zero temperature.

Next, we propagate this generalization to the energy functional to be minimized. This time, however, the energy function is to be minimized not only with respect to the orbitals but also with respect to the occupation numbers, which satisfy

$$\sum_{i=1}^{\infty} n_i = N \quad \text{with} \quad 0 \leq n_i \leq 1. \quad (51)$$

In order to facilitate the derivation, we employ a suitable parametrization of the occupation numbers (due to Dreizler and Gross)

$$n_i = \cos^2 \theta_i. \quad (52)$$

The generalized energy is now

$$\tilde{E}[\psi_1, \dots; \gamma_i, \dots] = \sum_{i=1}^{\infty} n_i t_i + \int n(\vec{r}) V_{eff}(\vec{r}) d\vec{r} \quad (53)$$

with the above definition of  $n(\vec{r})$  in Eq. 50. In Eq. 53, we have employed Janak's notation (modified slightly)

$$t_i = -\frac{1}{2} \int \psi_i^*(\vec{r}) \nabla^2 \psi_i(\vec{r}) d\vec{r} = \epsilon_i - \int \psi_i^*(\vec{r}) (V_H + V_{xc} + v_{ext}(\vec{r})) \psi_i(\vec{r}) d\vec{r} \quad (54)$$

where the latter equality follows straightforwardly from Eq. 46.

The generalized functional incorporating also the constraints that the orbitals are to remain orthonormal and the occupations should add up to  $N$ , in this case is (following the treatment of Dreizler and Gross)

$$Q[\psi_1, \dots; \theta_1, \dots] = \tilde{E}[\psi_1, \dots; \theta_1, \dots] - \sum_i \lambda_i \left[ \int |\psi_i|^2 d\vec{r} - 1 \right] - \mu \left[ \sum_{i=1}^{\infty} n_i - N \right] \quad (55)$$

where, as always, a diagonal matrix of Lagrange multipliers have been employed for the orthonormality condition. Differentiating  $Q$  with respect to the orbitals, we recover the Kohn-Sham equations with  $\lambda_i = n_i \epsilon_i$ . Differentiating on the other hand with respect to  $\theta_i$  gives us

$$(\sin 2\theta_i)[\langle \psi_i | \hat{h} | \psi_i \rangle - \mu] = (\sin 2\theta_i)[\epsilon_i - \mu] = 0 \quad (56)$$

where  $\hat{h}$  is the single-particle Kohn-Sham Hamiltonian. This equation yields three possibilities

$$\begin{aligned} \epsilon_i &= \mu \quad \text{and} \quad \theta_i \text{ is arbitrary} \Rightarrow 0 \leq n_i \leq 1 \\ \epsilon_i &\neq \mu \quad \text{and} \quad \theta_i \text{ is arbitrary} \Rightarrow n_i = 1 \\ \epsilon_i &\neq \mu \quad \text{and} \quad \theta_i \text{ is arbitrary} \Rightarrow n_i = 0 \end{aligned} \quad (57)$$

From Eq. 57, we get the same occupational pattern as the Kohn-Sham formalism except when the eigenvalue corresponds to the chemical potential, in which case the occupancy is allowed to be fractional. This then proves that  $Q[n]$  can be used to replace the usual Kohn-Sham energy functional and derivatives with respect to occupation numbers are permissible. Then the straightforward differentiation of  $\tilde{E}$  with respect to  $n_i$  in Eq. 53 gives

$$\frac{\partial \tilde{E}}{\partial n_i} = t_i + \int d\vec{r} V_{eff} |\psi_i|^2 = \epsilon_i, \quad (58)$$

which is the common statement of Janak's theorem. Note that in Janak's original work, the orbitals  $\psi_i$  have been taken to depend on the occupations in a self-consistent manner, so the treatment is different. A decade after Janak's theorem was put forward, its soundness was questioned by Valiev and Fernando in their 1995 paper (PRB, **52**, 10697). This work, however, didn't receive a lot of attention and Janak's theorem continues to be used even today. It seems to make practical sense in the same spooky fashion as the Kohn-Sham orbitals (see for instance Göransson *et. al.* PRB, 2005, **72**, 134204).

One important result that can be derived from Janak's theorem is the connection between the Kohn-Sham orbital energies and the energy differences between systems with  $N$  and  $N + 1$  particles. Since total energies can be differentiated in a continuous manner with respect to occupations (at least between systems with consecutive particle numbers), we can write

$$E_{N+1} - E_N = \int_N^{N+1} dn_M \frac{\partial E_M}{\partial n_M} = \int_N^{N+1} dn_M \epsilon_M \quad (59)$$

where  $\epsilon_M$  must, throughout the variation, remain the highest occupied orbital. Because of this Janak's theorem can only be applied to the extraction of an electron from or addition of an electron to the highest occupied orbital. This is in stark contrast to Koopmans' theorem for the Hartree-Fock treatment, where electrons can be removed from/added to any state. Once again though in practice, Janak's theorem is often used for arbitrary levels, in particular while calculating electron binding energies to core levels.

## I. PLANEWAVES

In the previous lecture, we arrived at the Kohn-Sham equation system which needs to be solved self-consistently. In shorthand notation, the Kohn-Sham system may be written as

$$\left[ \hat{T}_s + V_{eff} \right] \phi_i = \epsilon_i \phi_i \quad (1)$$

where  $\hat{T}_s$  is the single-particle kinetic energy,  $V_{eff}$  is the effective potential which contains the external, Hartree and exchange-correlation parts and  $\phi_i$  and  $\epsilon_i$  are the Kohn-Sham orbitals and eigenvalues. Just as for the Hartree-Fock equation system, it is very difficult if not impossible to solve the above equation in real space. Instead, a viable method is to expand each of the Kohn-Sham orbitals in a suitable basis, converting the real space equation in Eq. 1 into a matrix equation in coefficient space. A matrix equation is far easier to solve than a real-space equation.

A planewave expansion of the Kohn-Sham orbitals can be achieved as follows :

$$\phi_i(\vec{r}) = \sum_{\vec{q}} c_{i,\vec{q}} \frac{1}{\sqrt{\Omega}} \exp(i\vec{q} \cdot \vec{r}) \equiv \sum_{\vec{q}} c_{i,\vec{q}} |\vec{q}\rangle \quad (2)$$

where  $c_{i,\vec{q}}$  are the expansion coefficients and the normalization constant  $1/\sqrt{\Omega}$  is chosen so as to satisfy the definition of the Kronecker- $\delta$

$$\langle \vec{q}' | \vec{q} \rangle = \frac{1}{\Omega} \int_{\Omega} d\vec{r} \exp(-i\vec{q}' \cdot \vec{r}) \exp(i\vec{q} \cdot \vec{r}) = \delta_{\vec{q},\vec{q}'} \quad (3)$$

In order to convert Eq. 1 into a matrix equation, we insert Eq. 2 into Eq. 1, multiply from the left by  $\langle \vec{q}' |$  and integrate in real space

$$\sum_{\vec{q}} \langle \vec{q}' | \hat{H}_{eff} | \vec{q} \rangle c_{i,\vec{q}} = \epsilon_i \sum_{\vec{q}} \langle \vec{q}' | \vec{q} \rangle c_{i,\vec{q}} = \epsilon_i c_{i,\vec{q}'} \quad (4)$$

Eq. 5 is a matrix equation which may be written as

$$\bar{H}C = \epsilon_i C \quad (5)$$

where  $\bar{H}$  is the Hamiltonian in matrix representation and  $C$  is a vector of coefficients. In the planewave representation, the kinetic energy term assumes an extremely simple, diagonal form

$$\langle \vec{q}' | -\frac{1}{2} \nabla^2 | \vec{q} \rangle = \frac{1}{2} |\vec{q}|^2 \delta_{\vec{q},\vec{q}'} \quad (6)$$

For the potential part of the effective Hamiltonian, the Fourier transform is the natural way to proceed since the planewave basis is essentially the same object as the exponentials used while writing down the Fourier transform of operators. The effective potential,  $V_{eff}$  has the periodicity of the lattice and therefore the only allowed Fourier components are those with the wavevectors in the reciprocal space of the crystal. We thus have

$$V_{eff}(\vec{r}) = \sum_m V_{eff}(\vec{G}_m) \exp(i\vec{G}_m \cdot \vec{r}) \quad (7)$$

where

$$V_{eff}(\vec{G}) = \frac{1}{\Omega_{cell}} \int_{\Omega_{cell}} V_{eff}(\vec{r}) \exp(-i\vec{G} \cdot \vec{r}) d\vec{r} \quad (8)$$

and the index  $m$  has been attached to the reciprocal space vector so as to make the summation easier. If we now insert Eq. 8 into Eq. 4, we have

$$\langle \vec{q}' | V_{eff} | \vec{q} \rangle = \sum_m \langle \vec{q}' | V_{eff}(\vec{G}_m) | \vec{q} \rangle \exp(i\vec{G}_m \cdot \vec{r}) \quad (9)$$

$$\equiv \sum_m \langle \vec{q}' | V_{eff}(\vec{G}_m) | \vec{q} + \vec{G}_m \rangle = \sum_m V_{eff}(\vec{G}_m) \langle \vec{q}' | \vec{q} + \vec{G}_m \rangle = \sum_m V_{eff}(\vec{G}_m) \delta_{\vec{q}' - \vec{q}, \vec{G}_m} \quad (10)$$



Thus, the matrix elements of the effective potential are only nonzero for those wavevectors that differ by a reciprocal lattice vector. In order to switch completely to the representation in terms of the  $\vec{G}_m$ 's, we make the definitions  $\vec{q} = \vec{k} + \vec{G}_m$  and  $\vec{q}' = \vec{k} + \vec{G}_{m'}$ . Written this way,  $\vec{q}$  and  $\vec{q}'$  differ manifestly by a reciprocal lattice vector. The sum over  $\vec{q}$  in Eq. 4 is then converted into a sum over  $m$ . Then for a given  $\vec{k}$ , the Schrödinger-like equation is given by

$$\sum_m \langle \vec{k} + \vec{G}_{m'} | \hat{H}_{eff} | \vec{k} + \vec{G}_m \rangle c_{i,m} = \epsilon_i c_{i,m'} \quad (11)$$

In accord with the Bloch theorem, the  $\vec{k}$ -dependence in Eq. 11 may be incorporated into the Hamiltonian

$$\sum_m H_{m'm}(\vec{k}) c_{i,m}(\vec{k}) = \epsilon_i(\vec{k}) c_{i,m'}(\vec{k}) \quad (12)$$

where

$$H_{m'm}(\vec{k}) = \frac{1}{2} |\vec{k} + \vec{G}_m|^2 \delta_{m'm} + V_{eff}(\vec{G}_m - \vec{G}_{m'}) \quad (13)$$

Since the Hamiltonian is now labeled by  $\vec{k}$ , the coefficients and energies are also labeled by  $\vec{k}$ . As seen in Eq. 13, the effective potential depends on the difference between two reciprocal lattice vectors. This follows straightforwardly from the presence of the Kröneckers delta in its definition in Eq. 10. Next, let's look at the plane-wave representation of each term in the total energy (which will be useful later) and the potential one by one.

## II. HARTREE TERM IN PLANEWAVES

In order to find the Fourier representation of the Hartree term, we start by expanding the two densities in the definition as a Fourier sum

$$\begin{aligned} \frac{1}{2} \int \int d\vec{r} d\vec{r}' \frac{n(\vec{r}) n(\vec{r}')}{|\vec{r} - \vec{r}'|} &= \frac{1}{2} \sum_{\vec{G}, \vec{G}'} \int \int d\vec{r} d\vec{r}' \frac{n(\vec{G}) n(\vec{G}')}{|\vec{r} - \vec{r}'|} e^{-i\vec{G} \cdot \vec{r}} e^{-i\vec{G}' \cdot \vec{r}'} \\ &= \frac{1}{2} \sum_{\vec{G}, \vec{G}'} \int \int d\vec{u} d\vec{r}' \frac{n(\vec{G}) n(\vec{G}')}{|\vec{u}|} e^{-i\vec{G} \cdot \vec{u}} e^{-i(\vec{G} + \vec{G}') \cdot \vec{r}'} \\ &= \pi \sum_{\vec{G}, \vec{G}'} n(\vec{G}) n(\vec{G}') \int d\vec{r}' e^{-i(\vec{G} + \vec{G}') \cdot \vec{r}'} \int \int u^2 du (-d \cos \theta) \frac{e^{-iGu \cos \theta}}{u} \\ &= \pi \sum_{\vec{G}, \vec{G}'} n(\vec{G}) n(\vec{G}') \Omega_{cell} \delta_{\vec{G}\vec{G}'} \int u^2 \frac{e^{-iGu \cos \theta}}{-iGu^2} \Big|_1^{-1} du \\ &= \pi \Omega_{cell} \sum_{\vec{G}} n(\vec{G}) n(-\vec{G}) \int \frac{e^{iGu} - e^{-iGu}}{-iG} du \\ &= \pi \Omega_{cell} \sum_{\vec{G}} n(\vec{G}) n(-\vec{G}) \frac{e^{iGu} - e^{-iGu}}{G^2} \Big|_0^\infty \quad (14) \end{aligned}$$

$$\begin{aligned} &= \pi \Omega_{cell} \sum_{\vec{G}} n(\vec{G}) n(-\vec{G}) \lim_{\eta \rightarrow 0} \frac{e^{iG(u+i\eta)} - e^{-iG(u-i\eta)}}{G^2} \\ &= 2\pi \Omega_{cell} \sum_{\vec{G}} n(\vec{G}) n(-\vec{G}) \quad (15) \end{aligned}$$

where a *regulator* has been employed in order to take the complex integral. One final simplification comes from the fact that  $n(\vec{r}) = n(-v\vec{r})$ . This reflects in the Fourier components of  $n$  in the following manner :

$$n(\vec{r}) = \sum_{\vec{G}} n(\vec{G}) e^{-i\vec{G} \cdot \vec{r}} \quad (16)$$

On the other hand

$$n(-\vec{r}) = \sum_{\vec{G}} n(\vec{G}) e^{i\vec{G} \cdot \vec{r}} = \sum_{\vec{G}} n(-\vec{G}) e^{-i\vec{G} \cdot \vec{r}}. \quad (17)$$

Comparing Eq. 16 and Eq. 17, we obtain the following relation between the Fourier components

$$n(\vec{G}) = n(-\vec{G}) \quad (18)$$

Thus, the Fourier representation of the Hartree energy term is

$$\frac{1}{2} \int \int d\vec{r} d\vec{r}' \frac{n(\vec{r}) n(\vec{r}')}{|\vec{r} - \vec{r}'|} = 2\pi \Omega_{cell} \sum_{\vec{G}} \frac{n(\vec{G})^2}{G^2} \quad (19)$$

By a similar argument, the Hartree potential in the effective potential of Eq. 12 may be written in the Fourier representation as

$$V_H(\vec{G}) = 4\pi \frac{n(\vec{G})}{G^2} \quad (20)$$

Notice the absence of the volume factor in this case. This comes from the fact that there is one forward Fourier transform in this case which brings a factor of  $\frac{1}{\Omega}$  which is cancelled by the  $\Omega$  coming from the definition of the Krönecker delta.

### III. EXCHANGE-CORRELATION TERM IN PLANEWAVES

The exchange-correlation term is written in the form of an integral of the product of two functions. The Fourier transform then takes a particularly simple form, namely a convolution

$$\begin{aligned} \int d\vec{r} n(\vec{r}) \epsilon_{xc}(\vec{r}) &= \sum_{\vec{G}, \vec{G}'} \int d\vec{r} n(\vec{G}) e^{-i\vec{G} \cdot \vec{r}} \epsilon_{xc}(\vec{G}') e^{-i\vec{G}' \cdot \vec{r}} \\ &= \Omega_{cell} \sum_{\vec{G}, \vec{G}'} \delta_{\vec{G}, -\vec{G}'} n(\vec{G}) \epsilon_{xc}(\vec{G}') \\ &= \Omega_{cell} \sum_{\vec{G}} n(-\vec{G}) \epsilon(\vec{G}) = \Omega_{cell} \sum_{\vec{G}} n(\vec{G}) \epsilon_{xc}(\vec{G}) \end{aligned} \quad (21)$$

Following once again a similar argument for the exchange-correlation potential, we have

$$V_{xc}(\vec{G}) = \sum_{\vec{G}'} n_{xc}(\vec{G} - \vec{G}') \frac{d\epsilon_{xc}}{dn}(\vec{G}') + \epsilon_{xc}(\vec{G}) \quad (22)$$

### IV. THE EXTERNAL POTENTIAL — STRUCTURE AND FORM FACTORS

Let the crystal be composed of different species of atoms each labeled  $\kappa = 1, \dots, n_{sp}$ . For each  $\kappa$ , there are  $n_\kappa$  identical atoms at positions  $\vec{r}_{\kappa,j}$  in the unit cell. The ionic potential (the external potential) may then be written as a superposition of isolated atom potentials,

$$V_{ne}(\vec{r}) = \sum_{\kappa=1}^{n_{sp}} \sum_{j=1}^{n_\kappa} \sum_{\vec{T}} V^\kappa(\vec{r} - \vec{r}_{\kappa,j} - \vec{T}). \quad (23)$$

If we now consider the Fourier transform of such a function as in Eq. 23, this derivation follows :

$$\begin{aligned}
V(\vec{G}) &= \frac{1}{\Omega} \int_{\Omega} d\vec{r} V(\vec{r}) \exp(i\vec{G} \cdot \vec{r}) = \frac{1}{\Omega} \sum_{\kappa=1}^{n_{sp}} \sum_{j=1}^{n_{\kappa}} \int_{\Omega} d\vec{u} V^{\kappa}(\vec{u}) \exp(i\vec{G} \cdot (\vec{u} + \vec{\tau}_{\kappa,j})) \underbrace{\sum_{\vec{T}} \exp(i\vec{G} \cdot \vec{T})}_{N_{cell}} \\
&= \frac{1}{\Omega_{cell}} \sum_{\kappa=1}^{n_{sp}} \int_{\Omega} d\vec{u} V^{\kappa}(\vec{u}) \exp(i\vec{G} \cdot \vec{u}) \sum_{j=1}^{n_{\kappa}} \exp(i\vec{G} \cdot \vec{\tau}_{\kappa,j}) \\
&\equiv \sum_{\kappa=1}^{n_{sp}} \frac{\Omega^{\kappa}}{\Omega_{cell}} S^{\kappa}(\vec{G}) V^{\kappa}(\vec{G})
\end{aligned} \tag{24}$$

where

$$V^{\kappa}(\vec{G}) = \frac{1}{\Omega^{\kappa}} \int_{\Omega} V^{\kappa}(\vec{r}) \exp(i\vec{G} \cdot \vec{r}) d\vec{r} \tag{25}$$

is the form factor and

$$S^{\kappa}(\vec{G}) = \sum_j^{n_{\kappa}} \exp(i\vec{G} \cdot \vec{\tau}_{\kappa,j}) \tag{26}$$

is the structure factor. When separated in this way, the Fourier transform of the external potential reduces to the Fourier transform of the isolated atomic potential (to be approximated later by a pseudopotential) and a sum of planewaves which impose the periodicity.

Eq. 24 is particularly useful if the function is spherically symmetric since the Fourier transform of a spherically symmetric function is also spherically symmetric. Let  $f(\vec{r}) = f(r)$  for an arbitrary function,  $f$ . The Fourier transform is

$$\begin{aligned}
f(\vec{G}) &= -\frac{1}{\Omega} \int r^2 f(r) \exp(i\vec{G} \cdot \vec{r}) dr d\phi d\cos\theta \\
&= -\frac{2\pi}{\Omega} \int r^2 f(r) \frac{\exp(iGr \cos\theta)}{iGr} \Big|_0^{\pi} dr \\
&= -\frac{4\pi}{\Omega} \int r^2 f(r) \frac{\exp(-iGr) - \exp(iGr)}{iGr} dr \\
&= \frac{4\pi}{\Omega} \int r^2 j_0(|\vec{G}|r) f(r) dr \equiv f(|\vec{G}|)
\end{aligned} \tag{27}$$

We thus have

$$V^{\kappa}(\vec{G}) = V^{\kappa}(|\vec{G}|) \int_0^{\infty} dr r^2 j_0(|\vec{G}|r) V^{\kappa}(r) \tag{28}$$

## V. CALCULATION OF THE ELECTRON DENSITY

After the Kohn-Sham equations have been solved, the density needs to be calculated. Remember the fact that we have a different system of Kohn-Sham equations for each  $\vec{k}$ . For the final expression of the energy and the density, we have to conduct a summation of all these  $\vec{k}$  points over the first Brillouin zone. Thus, for the density,

$$n(\vec{r}) = \frac{1}{N_k} \sum_{\vec{k}, i} n_{i,\vec{k}} \quad \text{where} \quad n_{i,\vec{k}}(\vec{r}) = |\phi_{i,\vec{k}}| \tag{29}$$

where the index  $i$  runs over the occupied states. For the planewave basis, we obtain rather straightforwardly the following expression for  $n_{i,\vec{k}}(\vec{r})$

$$n_{i,\vec{k}} = \frac{1}{\Omega} \sum_{mm'} c_{im}^*(\vec{k}) c_{im'}(\vec{k}) e^{i(\vec{G}_{m'} - \vec{G}) \cdot \vec{r}} \tag{30}$$

which yields

$$n_{i\vec{k}}(\vec{G}) = \frac{1}{\Omega} \sum_m c_{i,m}^*(\vec{k}) c_{i,m''}(\vec{k}) \quad (31)$$

where  $m''$  is defined as the index of the vector  $\vec{G}_{m''} = \vec{G}_m + \vec{G}$ . Although it is a straightforward operation, obtaining the density through the use of coefficients of the Kohn-Sham orbitals requires  $N_{\vec{G}}^2$  operations. If, on the other hand, once we know the coefficients of the Kohn-Sham orbitals and convert them directly to a real space grid using efficient FFT algorithms, this only takes  $N_{\vec{G}} \log N_{\vec{G}}$  operations. If we know the real-space density, we can also compute  $V_{xc}$  and  $E_{xc}$  in real space easily. In summary, it is more advantageous to use  $n(\vec{G})$  to calculate the Hartree potential whereas for  $V_{xc}$  we use the real space density.

## VI. THE PLANEWAVE CUTOFF

If you have run a PW-DFT code before, you will have seen that there is a parameter in the input file that allows you to specify an energy cutoff. This cutoff is necessary because we cannot possibly have a complete expansion in terms of infinitely many planewaves. However, one can show that the contribution of the higher frequency planewaves are less than those of lower frequency planewaves. It is thus a viable option to introduce some sort of a cutoff and take only those wavevectors that satisfy

$$\frac{1}{2}|\vec{G}|^2 < E_{cut}. \quad (32)$$

What this cutoff should be depends upon the system and a convergence study should be made before deciding what's the appropriate cutoff.

We must keep in mind while talking about the efficiency of the FFT approach that the density requires a maximum Fourier component that has twice the frequency of that of the Kohn-Sham orbitals. Moreover, the FFT grid needs to be a regular parallelepiped whereas the  $E_{cut}$  in Eq. 32 defines a sphere. Thus, the FFT has some redundant kpoints which somewhat slow down the calculation, however it's still efficient enough to justify its use.

## Lecture VIII : The pseudopotential

### I. KOHN-SHAM PROBLEM FOR AN ISOLATED ATOM

For a one-electron atom, the Coulombic potential,  $V(\vec{r}) = V(r) = -Z/r$  is spherically symmetric. The solutions may then be split into a radial and an angular part

$$\psi_{nlm}(\vec{r}) = \psi_{nl}(r)Y_{lm}(\theta, \phi) = r^{-1}\phi_{nl}(r)Y_{lm}(\theta, \phi) \quad (1)$$

The last equality in Eq. 1 is given in order to simplify the spherically-symmetric Schrödinger equation and  $Y_{lm}(\theta, \phi)$  are normalized spherical harmonics. The spherically-symmetric Schrödinger equation for principle quantum number  $n$  can be reduced to the radial equation

$$-\frac{1}{2} \frac{d^2}{dr^2} \phi_{nl} + \left[ \frac{l(l+1)}{2r^2} + V_{ext}(r) - \epsilon_{n,l} \right] \phi_{n,l} = 0. \quad (2)$$

In the Kohn-Sham approach to the many-particle system, the form of the single-particle equations are identical to the above radial Schrödinger equation, with an effective potential,  $V_{eff}$  replacing the pure Coulomb potential. The substitution can be done directly provided  $V_{eff}$  is also spherically symmetric. Let's see if that is the case :

- Density :

$$n(\vec{r}) = \sum_{n,l,m}^{\text{occ}} |\psi_{n,l,m}(r)|^2 = \sum_{n,l}^{\text{occ}} (2l+1) |\psi_{n,l}(r)|^2 = n(r) \quad (3)$$

- External potential : Ionic Coulomb potential

$$V_{ext}(\vec{r}) = -Z/r = V_{ext}(r) \quad (4)$$

- The Hartree potential

$$\begin{aligned} V_H(\vec{r}) &= \int \frac{n(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r}' \\ &= \int \frac{n(r')}{[(r - r' \cos \theta \sin \phi)^2 + (r' \cos \theta \sin \theta)^2 + (r' \sin \theta)^2]^{1/2}} r'^2 dr' d(\cos \theta) d\phi = V_H(r) \end{aligned} \quad (5)$$

- The exchange-correlation potential

$$V_{xc}(\vec{r}) = \epsilon_{xc}[n(r)] + n(r) \frac{d\epsilon_{xc}}{dn}[n(r)] = V_{xc}(r) \quad (6)$$

- Total effective potential

$$V_{eff}(r) = V_{ext}(r) + V_H(r) + V_{xc}(r) \quad (7)$$

Thus the independent-particle Kohn-Sham states may be classified by the angular quantum numbers  $L = l, m$ . The one-particle equations are then analogous to the Schrödinger equation for the one-electron atom,

$$-\frac{1}{2} \frac{d^2}{dr^2} \phi_{nl} + \left[ \frac{l(l+1)}{2r^2} + V_{eff}(r) - \epsilon_{n,l} \right] \phi_{n,l} = 0. \quad (8)$$

## II. THE PSEUDOPOTENTIAL IDEA

The concept of a *pseudopotential* is related to replacing the effects of the core electrons with an effective potential. The pseudopotential generation procedure starts with the solution of the atomic problem using the Kohn-Sham approach. Once the KS orbitals are obtained, we make an arbitrary distinction between *valence* and *core* states. The core states are assumed to change very little due to changes in the environment so their effect is replaced by a model potential derived in the atomic configuration and it is assumed to be *transferable*. The valence states are seen to oscillate rapidly close to the core regions. With the introduction of the new potential, the valence states are made smoother.

Let's now work out the pseudopotential transformation in its most general framework : Assume that we have a Hamiltonian  $\hat{H}$ , core states  $\{|\chi_n\rangle\}$  and core eigenvalues  $\{E_n\}$ . We are looking at a single valence state  $|\psi\rangle$ . Let's replace the valence state by the smoother  $|\phi\rangle$  and expand the remaining portion in terms of core states :

$$|\psi\rangle = |\phi\rangle + \sum_n^{\text{core}} a_n |\chi_n\rangle \quad (9)$$

Next, we take the inner product of Eq. 9 with one of the core states. Because the valence state has to be orthogonal to the core states, we have

$$\langle\chi_n|\psi\rangle = \langle\chi_n|\phi\rangle + \underbrace{\sum_n^{\text{core}} a_n \langle\chi_m|\chi_n\rangle}_{a_m} = 0 \quad (10)$$

We now write the right-hand side of Eq. 10 in terms of the pseudofunction,  $|\phi\rangle$

$$|\psi\rangle = |\phi\rangle - \sum_n \langle\chi_n|\phi\rangle |\chi_n\rangle \quad (11)$$

Applying the Hamiltonian onto the expression in Eq. 11, we get

$$\hat{H}|\phi\rangle + \sum_n^{\text{core}} (E - E_n) |\chi_n\rangle \langle\chi_n|\phi\rangle = E|\phi\rangle \quad (12)$$

As a result, the smooth, pseudowavefunction satisfies an effective equation with the same eigenenergy of the real valence wavefunction. In the case of isolated atoms, the indices  $n$  corresponds to the combined index  $nlm_l$  including the principal quantum number  $n$ , the angular momentum quantum number  $l$  and the magnetic quantum number  $m_l$ . However, since the potential does not include any terms (such as  $\vec{L} \cdot \vec{S}$  coupling) that lift the  $m_l$  degeneracy, we include in what follows only  $n$  and  $l$ . The reason why the  $l$ -degeneracy is broken instead will become clear later on. Eq. 12 may then be written as an eigenvalue equation for the smooth *pseudo wavefunction* as

$$(\hat{H} + \hat{V}_{nl}) |\phi\rangle = E|\phi\rangle \quad (13)$$

where the extra potential  $\hat{V}_{nl}$  depends on  $l$  due to spherical symmetry and its effect is localized to the core. Since  $E > E_n$  by definition, it is a repulsive potential and it partially cancels the effect of the attractive Coulombic potential. The resulting potential is then a much weaker one than the original potential. This shows that the eigenstates of this new potential are smoother.

## III. NORM-CONSERVING PSEUDOPOTENTIALS

In order to ensure the optimum smoothness and transferability Hamann, Schlüter and Chiang (Phys Rev Lett **43** 1494 (1979)) came up with four criteria that the pseudopotentials are expected to obey in the *reference configuration* (i.e. the atomic configuration for which the pseudopotential is generated.) :

1. All-electron and pseudo eigenvalues agree for the reference configuration.

$$\hat{H}|\psi_{nl}^{AE}\rangle = \epsilon_{nl}|\psi_{nl}^{AE}\rangle \quad (14)$$

$$(\hat{H} + \hat{V}_{nl}) |\psi_{nl}^{PS}\rangle = \epsilon_{nl}|\psi_{nl}^{PS}\rangle \quad (15)$$

$$(16)$$

2. AE and PS wavefunctions agree beyond a certain cutoff,  $r_c$ .

$$\psi_{nl}^{AE}(r) = \psi_{nl}^{PS}(r) \quad \text{for } r \geq r_c \quad (17)$$

3. Real and pseudo norm squares integrated from 0 to  $R$  for all  $R < r_c$  agree.

$$\int_0^R |\phi_{nl}^{AE}|^2 r^2 dr = \int_0^R |\phi_{nl}^{PS}|^2 r^2 dr \quad (18)$$

4. The logarithmic derivatives and the energy derivative of the logarithmic derivative agree for all  $R < r_c$ .

$$\left[ (r\phi^{AE}(r))^2 \frac{d}{dE} \frac{d}{dr} \ln \phi^{AE}(r) \right]_R = \left[ (r\phi^{PS}(r))^2 \frac{d}{dE} \frac{d}{dr} \ln \phi^{PS}(r) \right]_R \quad (19)$$

Items 1 and 2 are obvious. Item 3 is necessary so that by Gauss' theorem, outside the *core region* where  $r < r_c$ , the electrostatic potential is identical for the real and pseudo wavefunctions. Finally, 4 is necessary in order to improve transferability, reproducing phase shifts of the real potential in the scattering sense. When we go from an atom to a different environment such as a solid or a molecule, the eigenvalues change. A pseudopotential satisfying this item reproduces these changes to linear order. Using the Friedel sum rule, one can show that item 4 is implied by item 3.

In order to generate a norm-conserving pseudopotential, we must follow the steps below :

1. Solve the all-electron atomic system.
2. Decide which of the states are to be taken as core and which are to be taken as valence.
3. Generate the pseudopotentials,  $V_{nl}(r)$ , for the valence states. There are several methods for doing that. We describe here the method due to Kerker. In this method, we assume a parametrized analytical function for the core region. The analytic form chosen by Kerker is  $\phi_{nl}^{PS}(r) = e^{p(r)}$ , where  $p(r)$  is a fourth-order polynomial matching the true solution and its various derivatives at  $r_c$ .
4. Once we have a form for the pseudowavefunction, invert the Schrodinger equation, yielding

$$V_{nl,total}(r) = \epsilon - \left[ \frac{l(l+1)}{2r^2} - \frac{\frac{d^2}{dr^2} \phi_{nl}^{PS}(r)}{\phi_{nl}^{PS}(r)} \right] \quad (20)$$

Keep in mind that we are dealing with  $\phi(r)$ , which is defined by  $\psi(r) = r\phi(r)$ .

5. The potential  $V_{l,total}(r)$  found by inversion of the Schrödinger equation is not only the ionic potential that we want, but includes also contributions from the Hartree and exchange-correlation terms. We must therefore subtract these two contributions.

$$V_{nl}(r) = V_{l,total}(r) - V_{Hxc}^{PS}(r) \quad (21)$$

where the Hartree and xc contributions are computed for the pseudowavefunctions. This operation is called *unscreening*.

There is no unique way to determine the  $V_{nl}$ 's. There are two opposing considerations :

1. Good transferability  $\Rightarrow$  small  $r_c$ .
2. Large  $r_c \Rightarrow$  smoother pseudopotentials.

A good pseudopotential is one that strikes a balance between these two constraints.

Although the generation procedure is similar, the form of the pseudopotential may be assumed differently depending on considerations such as ease of computation, accuracy etc. In the following subsections, we shall see two different ways of forming norm-conserving pseudopotentials. Keep in mind that these pseudopotentials are formed within the atomic configuration and then exported to the solid with the hope of decent transferability.

An actual example of the construction of the pseudopotential can be seen in Fig. 1 for the radial parts of 2S and 2P orbitals of the element B. The agreement of the wavefunctions after the cutoff radius  $r_c$  is evident.

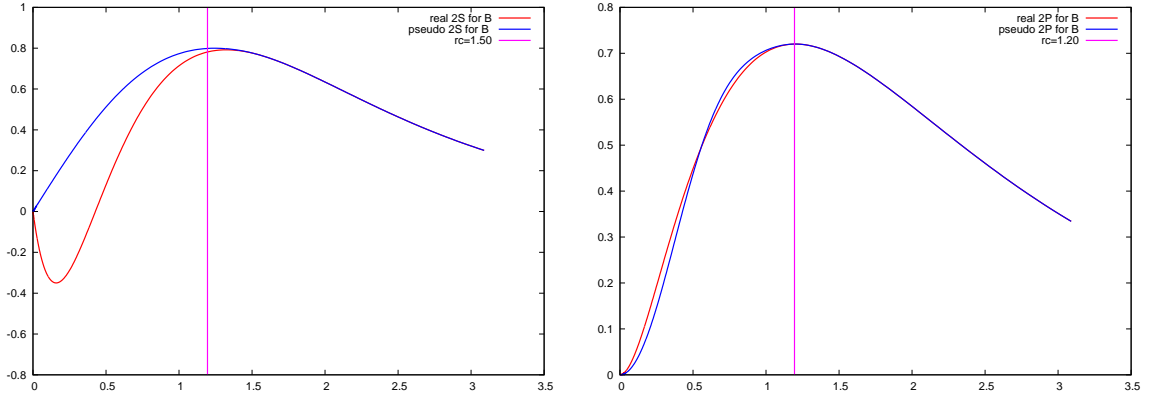


FIG. 1: Semilog plot of the real and pseudo wavefunctions for 2S (a) and 2P (b) orbitals of B.

### A. Semilocal pseudopotentials

Since a model pseudopotential replaces the potential of a nucleus, it is spherically symmetric and each angular momentum  $l$  may be treated separately. This leads to the *nonlocal*, i.e.  $l$ -dependent model pseudopotentials  $V_{nl}(r)$ . Thus, the general form of a pseudopotential may be written in the *semilocal* form as

$$\hat{V}_{SL} = \sum_{lm} |Y_{lm}\rangle V_{nl}(r) \langle Y_{lm}|. \quad (22)$$

Here the angular and the radial parts are separated due to spherical symmetry and the  $Y_{lm}$ 's are there because any function that depends upon angles only may be written as an expansion. This form is called semilocal because it is nonlocal only in the angular coordinates and it's local in the radial coordinate.

Although conceptually simple, semilocal pseudopotentials create a severe problem in the plane-wave basis representation due to the form of the nonlocal part. This can be seen by considering a particular matrix element of the pseudopotential and using a lot of straightforward but tedious algebra

$$\langle \vec{q} | \hat{V}_{SL} | \vec{q}' \rangle = \frac{1}{\Omega} \sum_{lm} \int \int e^{-i\vec{q} \cdot \vec{r}} Y_{lm}^*(\theta, \phi) V_{nl}(r) Y_{lm}(\theta, \phi) e^{i\vec{q}' \cdot \vec{r}'} r^2 d\Omega d\Omega' dr \quad (23)$$

where  $\vec{r}' = (r, \theta', \phi')$ . Notice the fact that we calculate a single integral over the radial component  $r$  but two integrals over the angles because the nonlocal part of the pseudopotential is in the form of *projectors* instead of a regular function as for the radial coordinate. In order to make use of their orthonormality, we expand the plane-waves in the integral in Eq. 23 in terms of spherical harmonics

$$e^{i\vec{q} \cdot \vec{r}} = 4\pi \sum_{lm} i^l j_l(qr) Y_{lm}^*(\hat{q}) Y_{lm}(\hat{r}) \quad (24)$$

where  $j_l(qr)$  are spherical Bessel functions and  $\hat{q}$  and  $\hat{r}$  denote the angles associated with the vectors  $\vec{q}$  and  $\vec{r}$ . Substituting Eq. 24 into Eq. 23 we have

$$\langle \vec{q} | \hat{V}_{SL} | \vec{q}' \rangle = \frac{(4\pi)^2}{\Omega} \sum_{lm} \sum_{l'm'} \sum_{l''m''} \int (-i)^{l'} i^{l''} j_{l'}(qr) j_{l''}(q'r) Y_{l'm'}(\hat{q}) Y_{l''m''}^*(\hat{q}') V_{nl}(r) r^2 dr \times \underbrace{\int Y_{l'm'}^*(\hat{r}') Y_{lm}(\hat{r}') d\Omega'}_{\delta_{l'l', mm'}} \underbrace{\int Y_{lm}^*(\hat{r}) Y_{l''m''}(\hat{r}) d\Omega}_{\delta_{l'l'', mm''}} \quad (25)$$

where we have used the orthonormality of spherical harmonics as mentioned before. When the delta functions are imposed the two sums over  $l'm'$  and  $l''m''$  are gone and the only indices that remain are  $l$  and  $m$ . Next in order to deal with the remaining spherical harmonics, we utilize the *summation theorem for spherical harmonics*.

$$\sum_m Y_{lm}(\hat{q}) Y_{lm}^*(\hat{q}') = \frac{2l+1}{4\pi} P_l(\cos \theta_{\vec{q}\vec{q}'} ) \quad (26)$$



where  $P_l$  is the generalized Legendre polynomial and  $\theta_{\vec{q}\vec{q}'}$  is the angle between the two wavevectors  $\vec{q}$  and  $\vec{q}'$ . Putting everything together the expectation value in Eq. 25 becomes

$$\langle \vec{q} | \hat{V}_{NL} | \vec{q}' \rangle = \frac{4\pi}{\Omega} \sum_l (2l+1) \int j_l(qr) j_l(q'r) P_l(\cos \theta_{\vec{q}\vec{q}'}) V_l(r) r^2 dr \quad (27)$$

As seen in Eq. 27, the semilocal pseudopotential mixes planewaves with different  $\vec{q}$  vectors. This means that if we have  $N_{pw}$  planewaves in the basis that we choose we have to evaluate  $N_{pw}^2$  such integrals as in Eq. 27, which brings a large computational load. In their seminal work Kleinman and Bylander (*Physical Review Letters*, **48**, 1425) proposed a different scheme for composing the angular momentum-dependent part of the pseudopotential such that the number of integrals to be calculated could be reduced from  $N_{pw}^2$  to just  $N_{pw}$ . This form is called the *fully nonlocal form* and we shall see how it is formed in the next section.

### B. The Kleinman-Bylander Transformation

The difficulty with the semilocal pseudopotentials above is that due to the locality in the radial coordinate there is a single radial variable for both planewaves while the angular part is nonlocal and the integrals related to each angular coordinate may be evaluated separately. The fully nonlocal formulation includes the radial part in the projectors thereby separating the integral related to each projector. The starting point for the transformation is the usual semilocal form explained above where the local radial part is given by  $V_l$ . We then find the pseudowavefunctions in the atomic configuration,  $\phi_l^{PS}$  corresponding to this pseudopotential. Using these wavefunctions, we construct a new pseudopotential who is composed of a local part,  $V_L$ , which is a completely arbitrary function that has the long-range behavior of the real potential and a nonlocal part  $V_{NL}$ , which is the remaining part of the pseudopotential. After the subtraction of the local part, the nonlocal part is left behind localized in the core region. The nonlocal part is then given by

$$\hat{V}_{NL} = \sum_l \hat{V}_l^{NL} = \sum_{lm} \frac{|\hat{V}_l^{SL} \phi_{lm}^{PS}\rangle \langle \phi_{lm}^{PS} | \hat{V}_l^{SL}|}{\langle \phi_{lm}^{PS} | \hat{V}_l | \phi_{lm}^{PS} \rangle} \quad (28)$$

This particular form is chosen such that its action on a pseudowavefunction yields the same result as the action of the semilocal form

$$\hat{V}_l^{NL} |\phi_{lm}^{PS}\rangle = \hat{V}_l^{SL} |\phi_{lm}^{PS}\rangle \quad (29)$$

In order to see this, let's apply this operator on the pseudowavefunction

$$\hat{V}_l^{NL} |\phi_{lm}^{PS}\rangle = \frac{|\hat{V}_l^{SL} \phi_{lm}^{PS}\rangle \langle \phi_{lm}^{PS} | \hat{V}_l^{SL} | \phi_{lm}^{PS} \rangle}{\langle \phi_{lm}^{PS} | \hat{V}_l^{SL} | \phi_{lm}^{PS} \rangle} = \hat{V}_l^{SL} |\phi_{lm}^{PS}\rangle \quad (30)$$

In the planewave representation, an arbitrary matrix element of the Kleinman-Bylanderized pseudopotential is

$$\langle \vec{q} | \hat{V}_l^{NL} | \vec{q}' \rangle = \frac{\langle \vec{q} | \hat{V}_l \phi_{lm}^{PS} \rangle \langle \hat{V}_l \phi_{lm}^{PS} | \vec{q}' \rangle}{\langle \phi_{lm}^{PS} | \hat{V}_l | \phi_{lm}^{PS} \rangle}. \quad (31)$$

The two integrals in Eq. 31 do not share any common integral variables and may thus be performed separately. One could then compute  $N_{PW}$  such integrals, i.e.  $\langle \vec{q} | \hat{V}_l \phi_{lm}^{PS} \rangle$  in advance and calculate the appropriate product  $\langle \vec{q} | \hat{V}_l \phi_{lm}^{PS} \rangle \langle \hat{V}_l \phi_{lm}^{PS} | \vec{q}' \rangle$  for each  $\vec{q}, \vec{q}'$  pair.

## IV. ULTRASOFT PSEUDOPOTENTIALS

In 1990, Vanderbilt (see *Physical Review B*, **41**, 7892) proposed a new and radical method for generating pseudopotentials by relaxing the norm-conservation constraint. Pseudopotentials generated in this way (due to their softness) require a much smaller planewave cutoff and thus a much smaller number of planewaves. For this reason they are usually called *ultrasoft pseudopotentials*.

First let's see how ultrasoft pseudopotential is integrated into the Kohn-Sham equations and then see how it is generated. The total energy in the presence of ultrasoft pseudopotentials is written as

$$E_e = \sum_i \sum_i \langle \phi_i | -\frac{1}{2} \nabla^2 + V_{NL} | \phi_i \rangle + \frac{1}{2} \int \int d\vec{r} d\vec{r}' \frac{n(\vec{r})n(\vec{r}')}{|\vec{r} - \vec{r}'|} + E_{xc}[n] + \int d\vec{r} V_{loc}(\vec{r})n(\vec{r}) \quad (32)$$

where the Hartree, exchange-correlation, local and kinetic terms have the usual form. The nonlocal part of the pseudopotential instead is given in terms of a new set of projection operators as

$$V_{NL} = D_{nm}^{(0)} |\beta_n^I\rangle \langle \beta_m^I| \quad (33)$$

where the superscript  $I$  indicates that the projector is centered on atom  $I$ .

The norm-conserving condition is relaxed by introducing a generalized orthonormality condition through the overlap matrix  $S$

$$\langle \phi_i | S | \phi_j \rangle = \delta_{ij} \quad (34)$$

where

$$S = 1 + \sum_{nm,I} q_{nm} |\beta_n^I\rangle \langle \beta_m^I| \quad (35)$$

with  $q_{nm} = \int d\vec{r} Q_{nm}(\vec{r})$ . Because the norm-conserving condition is now relaxed, the electronic density needs to be *augmented* in order to preserve the correct number of electrons. The new density is then

$$n(\vec{r}) = \sum_i \left[ |\phi_i(\vec{r})|^2 + \sum_{nm,I} Q_{nm}^I(\vec{r}) \langle \phi_i | \beta_n^I \rangle \langle \beta_m^I | \phi_i \rangle \right] \quad (36)$$

In order to verify that the density as given in Eq. 36 integrates to the correct number of electrons, consider

$$\begin{aligned} \int n(\vec{r}) d\vec{r} &= \sum_i \left[ \int |\phi_i(\vec{r})|^2 + \sum_{nm,I} \int d\vec{r} Q_{nm}^I(\vec{r}) \langle \phi_i | \beta_n^I \rangle \langle \beta_m^I | \phi_i \rangle \right] \\ &= \sum_{ij} \delta_{ij} \left[ \langle \phi_i | \phi_j \rangle + \sum_{nm,I,j} \delta_{ij} q_{nm} \langle \phi_i | \beta_n^I \rangle \langle \beta_m^I | \phi_j \rangle \right] \\ &\equiv \sum_{ij} \delta_{ij} \langle \phi_i | S | \phi_j \rangle = \sum_{ij} \delta_{ij} \delta_{ij} = \sum_i \delta_{ii} = \sum_i 1 = N_v \end{aligned} \quad (37)$$

where  $N_v$  stands for the number of valence electrons in the system.

The new form of the density, although it gives us the flexibility to create softer pseudopotentials, inevitably makes the Kohn-Sham equations more complex. The functional derivative of the density with respect to the orbitals is

$$\frac{\delta n(\vec{r}')}{\delta \phi_i^*(\vec{r})} = \phi_i(\vec{r}) \delta(\vec{r} - \vec{r}') + \sum_{nm,I} Q_{nm}^I(\vec{r}') \beta_n^I(\vec{r}) \langle \beta_m^I | \phi_i \rangle. \quad (38)$$

Utilizing Eq. 38 and Eq. 32 we arrive at the modified Kohn-Sham equations

$$\frac{\delta E_e}{\delta \phi_i^*} = \int d\vec{r}' \frac{\delta E_e}{\delta n(\vec{r}')} \frac{\delta n(\vec{r}')}{\delta \phi_i^*(\vec{r})} = \left[ -\frac{1}{2} \nabla^2 + V_{eff} + \left( D_{nm}^{(0)} + \int d\vec{r}' V_{eff}(\vec{r}') Q_{nm}(\vec{r}') \right) |\beta_n^I\rangle \langle \beta_m^I| \right] |\phi_i\rangle. \quad (39)$$

where  $V_{eff}$  is the regular potential

$$V_{eff} = V_H + V_{loc} + V_{xc}. \quad (40)$$

As easily seen in Eq. 39, unlike norm-conserving pseudopotentials, the coefficients of the projectors in the nonlocal part of the pseudopotential gets updated in a self-consistent manner at every iteration. This introduces a slight overhead to the calculation but the gain in the planewave expansion surpasses the loss thus introduced.

Now, we take a step back and see the justification and generation of such a radical implementation. The generation of ultrasoft potentials start out in the same way as norm-conserving pseudopotentials. For a given species, the atomic

Kohn-Sham system is solved self-consistently, resulting in the screened all-electron potential,  $V_{AE}$ . Then for each angular momentum channel, a few energy values,  $\epsilon_{l\tau}$  are chosen ( $\tau$ =number of such energies, usually max 3). For the  $V_{AE}$  already obtained, the Schrödinger equation is found

$$[T + V_{AE}]|\psi_n(r)\rangle = \epsilon_n|\psi_n(r)\rangle \quad (41)$$

where  $n$  is a composite index holding  $l$  for the angular momentum quantum number,  $m$  for the projection and  $\tau$ . In the general case,  $\psi_n$  is not going to be normalizable but we use the bra-ket notation anyway.

Next, we choose *three* cutoff radii :

1. A smooth local ( $l$ -independent) potential is created which matches  $V_{AE}(r)$  after the radius  $r_c^{loc}$ .
2. A second cutoff radius,  $r_{cl}$ , is chosen for each angular momentum channel and a pseudowavefunction,  $\phi_n$  is constructed which matches the real wavefunction  $\psi_n$  at  $r_{cl}$ .
3. Finally a diagnostic radius  $R$  is chosen which is slightly larger than the maximum of the maximum of all the  $r_{cl}$  and  $r_C^{loc}$ .

Then new orbitals are formed that satisfy

$$|\chi_n\rangle = (\epsilon_n - T - V_{loc})|\phi_n\rangle \quad (42)$$

which are zero beyond  $R$ .

Next, an auxiliary matrix of inner products

$$B_{nm} = \langle \phi_n | \chi_m \rangle \quad (43)$$

and the projectors necessary for the definition of the nonlocal part of the potential

$$|\beta_n\rangle = \sum_m (B^{-1})_{mn} |\chi_m\rangle \quad (44)$$

are defined.

While determining the pseudowavefunctions mentioned above, norm-conservation has not been imposed. We thus define a set of generalized *augmentation functions*

$$Q_{nm}(\vec{r}) = \psi_n^*(\vec{r})\psi_m(\vec{r}) - \phi_n^*(\vec{r})\phi_m(\vec{r}) \quad (45)$$

and the related *augmentation charges*

$$q_{nm} = \langle \psi_n | \psi_m \rangle_R - \langle \phi_n | \phi_m \rangle_R. \quad (46)$$

where the subscript  $R$  denotes an integration within the cutoff radius,  $R$ . With all these descriptions, we can verify that the pseudowavefunctions  $|\phi_n\rangle$  obey the equation

$$\left( \hat{T} + V_{loc} + \sum_{nm} D_{nm} |\beta_n\rangle \langle \beta_m| \right) |\phi_i\rangle = \epsilon_i \left( 1 + \sum_{nm} q_{nm} |\beta_n\rangle \langle \beta_m| \right) |\phi_i\rangle \quad (47)$$

where  $D_{nm} = B_{nm} + \epsilon_m q_{nm}$ . In order to prove the central equation Eq. 47, let's look at

$$\begin{aligned} \left( \sum_{nm} B_{nm} |\beta_n\rangle \langle \beta_m| \right) |\phi_i\rangle &= \sum_{nm} B_{nm} \sum_l (B^{-1})_{ln} |\chi_l\rangle \sum_k (B_{km}^{-1})^* \underbrace{\langle \chi_k | \phi_i \rangle}_{B_{ik}^*} \\ &= \sum_{ml} |\chi_l\rangle \sum_n (B^{-1})_{ln} B_{nm} \sum_k (B_{ik} (B^{-1})_{km}) \\ &= \sum_{ml} \delta_{lm} \delta_{im} |\chi_l\rangle = |\chi_i\rangle = (\epsilon_i - \hat{T} - V_{loc}) |\phi_i\rangle \end{aligned} \quad (48)$$

Sustituting Eq. 48 into Eq. 47 completes the proof.

Thus the fundamental equation of the pseudopotential theory has been satisfied, that is pseudowavefunctions that satisfy the same equation as the real wavefunctions have been found. These pseudopotentials, however, may be proven to satisfy, instead of being orthonormal in the usual sense, satisfy a generalized orthonormality equation through an overlap matrix

$$S = 1 + \sum_{nm} q_{nm} |\beta_n\rangle \langle \beta_m| \quad (49)$$

Once the  $D_{nm}$  are obtained, the local part and the nonlocal coefficients of the bare pseudopotential are obtained through a descreening procedure

$$V_{loc}^{ion} = V_{loc} - V_H - V_{xc} \quad (50)$$

$$D_{nm}^{(0)} = D_{nm} - \int d\mathbf{r}' V_{loc}(\mathbf{r}') n(\mathbf{r}') \quad (51)$$

While generating pseudopotentials, we don't enforce norm conservation. This allows us to increase the cutoff radii to larger values. This convenient fact however brings about the necessity of defining a separate planewave cutoff for the density because as seen in the expression in Eq. 36, the density now has a smooth part and a sharp part. This sharp part needs to be expressed using planewaves that have high frequencies, i.e. a plane wave cutoff that is typically several times the planewave cutoff used to express the wavefunctions.

## I. INTEGRATION

We have seen that in crystals the Kohn-Sham eigenvalues need to be found which are self-consistent solutions of a  $\vec{k}$ -dependent Hamiltonian

$$H(\vec{k})\phi_{\vec{k}}(\vec{r}) = \epsilon_i(\vec{k})\phi_{\vec{k}}(\vec{r}). \quad (1)$$

The total energy is then given through an integration over the first Brillouin zone. In numerical practice we convert this integral into a sum over discrete points in  $\vec{k}$ -space. We then have

$$E_e = \frac{1}{N_k} \sum_{\vec{k}, i} \left[ \epsilon_i(\vec{k}) + \sum_{\vec{G}} \epsilon_{xc}(\vec{G}) n(\vec{G}) + 2\pi \sum_{\vec{G}} \frac{n(\vec{G})^2}{G^2} \right] \quad (2)$$

In principle if we manage to choose a dense enough  $\vec{k}$ -grid over which to perform the summation, it makes little difference how many  $\vec{k}$ -points we have in our sum or the manner in which they are chosen. However, since each new  $\vec{k}$ -point means a new self-consistent Kohn-Sham system it's crucial to maximize the accuracy while reducing the size of the  $\vec{k}$ -point set as much as possible.

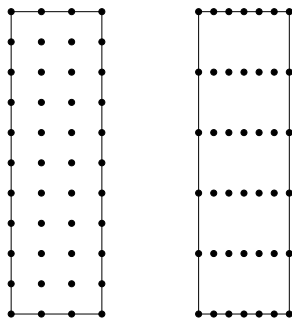
Since all functions having the symmetry of the crystal (such as the density) repeat beyond the first Brillouin zone, the  $\vec{k}$ -points are also constrained to this area. According to a widely-used procedure proposed by Monkhorst and Pack (see *Physical Review B*, **13**, 5188) a uniform set of points may be determined using the formula

$$\vec{k}_{n_1, n_2, n_3} = \sum_i^3 \frac{2n_i - N_i - 1}{2N_i} \vec{b}_i \quad (3)$$

where  $\vec{G}_i$  are the primitive reciprocal lattice vectors,  $n_i = 1, \dots, N_i$  and  $N_i$  determines the number of  $\vec{k}$ -points to be chosen in each direction.

### A. Choosing kpoints

There are several considerations while choosing kpoints.



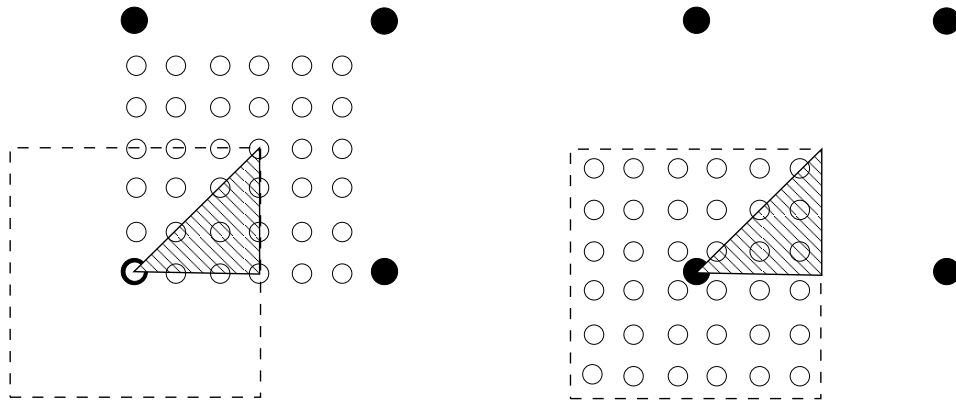
### B. Uniformity

As discussed in the illuminating paper by Moreno and Soler (*Physical Review B*, **45**, 24), mesh points that are uniformly distributed are preferred over those that have different densities in different dimensions. Let's assume that we have a crystal that whose BZ is longer in one dimension than the other as seen in the figure to the left. Although the two meshes illustrated in the figure have more or less the same number of points, the one on the right resolves even small oscillation in one direction while leaving the other direction inadequately sampled. In the one on the left, however, the resolution in either direction is equally uniform and thus this is preferred over

the one on the right since there is no reason to expect oscillations particularly in one direction.

### C. Symmetry

While summing a function over kpoints in the BZ, it reduces the calculation time by a several factors if we make use of the symmetry of the crystal in question. Points in the reciprocal space grid generated for the BZ integration are usually connected to each other by symmetry operations. Using these operations, the points to be used in the integration may be reduced to what is known as the *irreducible BZ*.



Integrals over the BZ may be interchanged with integrals over the IBZ

$$\bar{f}_i = \sum_{\vec{k}}^{IBZ} w_{\vec{k}} f_i(\vec{k}) \quad (4)$$

where  $w_{\vec{k}}$  are the relative weights.

## II. INSULATORS VS METALS

What differentiates between an insulator and a metal is the presence of a gap and the location of the Fermi level. The Fermi level is the energy of the highest occupied state and if it falls inside a gap in the band structure, then the crystal is said to be an insulator. If on the other hand it falls on a point where there is a finite density of states, it is a conductor. In terms of  $\vec{k}$ -integration, the insulators are well-behaved in the sense that the density of states goes to zero smoothly before the gap and the integration of a smooth function usually does not cause problems. For metals on the other hand, the functions to be integrated need first to be multiplied by a sharp Fermi occupation, which makes their resolution in planewaves very hard.

$$I = \int_{BZ} f(\vec{k}) [1 - \theta(E(\vec{k}) - E_F)] d\vec{k} \quad (5)$$

In order to overcome the difficulty we introduce an artificial smearing to the problem. There are several schemes for this smearing, each with their advantages and disadvantages. First we recast the integral into a form with a delta-function (as suggested in the paper by Methfessel and Paxton *Physical Review B*, **40**, 3616) which yields for Eq. 5

$$I = \int_{-\infty}^{\infty} [1 - \theta(E(\vec{k}) - E_F)] \int_{BZ} f(\vec{k}) \delta(\epsilon - E(\vec{k})) d\epsilon \quad (6)$$

Next, we replace the delta function by a smearing function which makes the integrand smoother. Several methods have been proposed for this smoothing procedure. Each method, while improving the  $\vec{k}$ -point convergence, causes the answer to converge to the wrong value. Among the methods available we seek those that minimize this error.

- **Fermi-Dirac function :** This introduces an artificial temperature into the system by approximating the step function by a Fermi-Dirac distribution

$$f(E) = \frac{1}{e^{(E-E_F)/kT} + 1} \quad (7)$$

which as  $T \rightarrow 0$  approaches the step-function. This approach has a drawback however that the reduced occupation probability in the states with energies smaller than the Fermi energy are not compensated by the occupation probability of the newly introduced states with energies higher than the Fermi energy.

- **Methfessel-Paxton** : The Methfessel-Paxton scheme aims to overcome this shortcoming by introducing a smearing scheme that can progressively be improved. They first propose an approximation scheme to the delta-function and then integrate it to obtain an approximation to the step function. The form they propose is an expansion in terms of *Hermite functions* which are simply Hermite polynomials multiplied by Gaussians.

$$\delta(x) \approx D_N = \sum_n^N A_n H_{2n} e^{-x^2} \quad (8)$$

where  $A_n$  are suitable coefficients. They use Hermite polynomials of even order only since the delta function is even. The delta function obtained by integrating  $D_N$  in the following way

$$\delta(x) \approx S_N = 1 - \int_{-\infty}^x D_N(t) dt \quad (9)$$

Using recursion properties of Hermite functions, we arrive at

$$S_0 = \frac{1}{2}(1 - \text{erf}(x)) \quad (10)$$

$$S_N = S_0(x) + \sum_{n=1}^N A_n H_{2n-1}(x) e^{-x^2} \quad (11)$$

$$(12)$$

Here  $S_0$  corresponds to the Fermi-Dirac smearing whereas the other terms serve to correct errors introduced.

A drawback of the Methfessel-Paxton smearing is that we encounter negative occupation values which are unrealistic and cause problems in such procedures as visualizing the density of states.

- **Gaussian** : Approximate delta function by a Gaussian obtain delta-function by integration.
- **Marzari-Vanderbilt** : The Marzari-Vanderbilt approach has been proposed in order to amend the negative occupancies introduced by the Methfessel-Paxton method. The delta-function this time is approximated by a Gaussian multiplied by a first order polynomial

$$\tilde{\delta}(x) = \frac{1}{\sqrt{\pi}} e^{-[x-(1/\sqrt{2})]^2} (2 - \sqrt{2}x) \quad (13)$$

where

$$x = \frac{\mu - \epsilon}{\sigma} \quad (14)$$

with width  $\sigma$ .

## Lecture X : Self-consistent solution of the Kohn-Sham equation

### I. SELF-CONSISTENT NATURE OF THE DENSITY IN THE KOHN-SHAM SYSTEM

In this lecture and the following, the topics that we shall discuss are going to be technical rather than physical. The methods that we shall look at do not bear relevance as far as density functional theory is concerned but they are central to the efficient execution of the calculations.

As shown in previous lectures, the Kohn-Sham equations have each the form of a single-particle Schrödinger equation with the difference that the effective potential operator, which yields the orbitals that determine the density, depends on the density itself. The Kohn-Sham system then must be solved self-consistently. This is to say that taking as input an initial density, a new density is found. This new density is then fed into the equation system which yields an output density and so on. This simplistic approach however usually results in poor convergence properties. In general, we resort to a set of algorithm which are called *mixing algorithms*. These carefully designed algorithms prescribe ways of updating the density by mixing old and new densities in the best possible way. In this lecture, we shall see why the simple-minded approximation fails to converge and how the more sophisticated algorithms work. For a very accessible description of the problem see the paper by Dederichs and Zeller (*Physical Review B*, **28**, 5462, 1983).

At the  $N$ th step of the self-consistent iterations, the output density (which is to be read in in the  $N + 1$ th step) may be expressed as a functional of the input density, i.e.

$$n_{N+1} = F[n_N] \quad (1)$$

Let's now assume that we are rather close to the minimum. This assumption is necessary for the rest of the treatment since we are going to assume linear response. To linear order, the difference between the current density and ground state density (the one we are trying to reach) denoted by  $\delta n(\vec{r}) = n(\vec{r}) - n_0(\vec{r})$  may be written in terms of the difference in the potential

$$\delta n(\vec{r}) = \int d\vec{r}' \chi(\vec{r}, \vec{r}') \delta V(\vec{r}'). \quad (2)$$

The response function,  $\chi(\vec{r}, \vec{r}')$  in Eq. 2 may be shown to be related to the dielectric tensor of the system. Due to the self-consistent nature of the system on the other hand, the potential may just as well be regarded as the independent variable and the difference in the potential may be expressed in terms of the difference in the density

$$\delta V(\vec{r}) = \int d\vec{r}' U(\vec{r}, \vec{r}') \delta n(\vec{r}') \quad (3)$$

where by definition  $U(\vec{r}, \vec{r}')$  is the functional derivative of the potential  $V(\vec{r})$ . In other words it is the second functional derivative of the total energy. It can easily be evaluated to yield

$$U(\vec{r}, \vec{r}') = \frac{1}{|\vec{r} - \vec{r}'|} + \left. \frac{d\epsilon_{xc}}{dn} \right|_{n=n_0} \quad (4)$$

Substituting Eq. 3 into Eq. 2, we obtain (paying careful attention to the integration variables)

$$\begin{aligned} \delta \rho(\vec{r}) &= \int d\vec{r}' \int d\vec{r}'' \chi(\vec{r}, \vec{r}') U(\vec{r}', \vec{r}'') \delta n(\vec{r}'') \\ &= \int d\vec{r}'' \underbrace{\int d\vec{r}' \chi(\vec{r}, \vec{r}') U(\vec{r}', \vec{r}'')}_{f(\vec{r}, \vec{r}'')} \delta n(\vec{r}'') \\ &= \int d\vec{r}' f(\vec{r}, \vec{r}') n(\vec{r}') \end{aligned} \quad (5)$$

where to obtain the last line dummy variables have been exchanged. Eq. 5 displays plainly the self-consistent nature of the density. The iterative version of Eq. 5 is

$$\delta n_{N+1} = \underline{f} \delta n_N \quad (6)$$

where the integral operator in Eq. 5 has been denoted by the symbol  $\underline{f}$ . By recursive application of Eq. 6, we arrive at

$$\delta n_{N+1} = \underline{f} n_N = \underline{f}^N \delta n_1 \quad (7)$$



Just like any operator,  $\underline{f}$  can be written in the spectral representation with eigenvalues  $\{\lambda_i\}$  and eigenvectors  $\{|i\rangle\}$ . Then Eq. 7 becomes

$$\delta n_{N+1} = \sum_i \lambda_i^N |i\rangle\langle i| \delta n_1 \quad (8)$$

As is evident from the form in Eq. 8 convergence is achieved in this simple-minded approach only if all the eigenvalues of the operator  $\underline{f}$  is between -1 and 1. Since this cannot be guaranteed in general such an iterative scheme diverges for most (if not all) systems and a mixing scheme must be applied.

## II. MIXING SCHEMES

### A. Simple mixing

The simplest algorithm used for mixing is the addition of a certain fraction of the density in the previous step to that of the current step, that is to say

$$n_{N+1}(\vec{r}) = \alpha F[n_N] + (1 - \alpha)n_N \quad (9)$$

where  $F$ , defined in Eq. 1 is the functional which yields the output density given an input density and  $\alpha$  is a constant determined at the beginning of the calculation. In terms of the linear differences Eq. 9 may be written as

$$\begin{aligned} \delta n_{N+1} &= \alpha \underline{f} \delta n_N + (1 - \alpha) \delta n_N = [1 - \alpha(1 - \underline{f})] \delta n_N \\ &= \sum_i (1 - \alpha(1 - \lambda_i))^N |i\rangle\langle i| \delta n_1 \equiv \sum_i (1 - \alpha\mu_i)^N |i\rangle\langle i| \delta n_1 \end{aligned} \quad (10)$$

where the spectral representation of the operator has once again been employed and the definition  $\mu_i \equiv 1 - \lambda_i$  has been made.

For convergence the inequality  $|1 - \alpha\mu_i|$  must be satisfied. In other words if we manage to choose  $\alpha$  in the *convergence region*

$$0 < \alpha < 2/\mu_{\max} \quad (11)$$

then converge is achieved in the self-consistent iterations. The value  $\alpha$  needs to be chosen differently for different systems. While certain systems require very small  $\alpha$  (meaning slow convergence), other systems are more stable and converge more rapidly with larger values of  $\alpha$ .

### B. Anderson mixing

Anderson mixing is an accelerated mixing scheme where both the input and the output densities are mixed with suitable mixing parameters.

$$\begin{aligned} \tilde{n}_N &= \alpha_N n_N + (1 - \alpha_N) n_{N-1} \\ \tilde{F}_N &= \alpha_N F[n_N] + (1 - \alpha_N) F[n_{N-1}] \end{aligned} \quad (12)$$

In contrast to the simple mixing scheme illustrated in Eq. 10,  $\alpha$  is not taken to be a constant for the entire duration of the calculation but updated at each iteration step in such a way that

$$\int d\vec{r} [\tilde{F}_N(\vec{r}) - \tilde{n}_N(\vec{r})]^2 \quad (13)$$

is minimized. This is the second best thing to use an  $\vec{r}$ -dependent  $\alpha$ . Defining

$$r_n \equiv F_N - f_N \quad \text{and} \quad \int f(\vec{r}) g(\vec{r}) d\vec{r} \equiv (f, g) \quad (14)$$

we expand out the terms in Eq. 13 using Eq. 12 to get

$$\begin{aligned} \int d\vec{r} [\tilde{F}_N(\vec{r}) - \tilde{n}_N(\vec{r})]^2 &= \int d\vec{r} [\alpha_N^2 (F_N - n_N)^2 + (1 - \alpha_N)^2 (F_{N-1} - n_{N-1})^2 + 2\alpha_N(1 - \alpha_N)(F_N - n_N)(F_{N-1} - n_{N-1})] \\ &= \alpha_N^2 (r_N, r_N) + (1 - \alpha_N)^2 (r_{N-1}, r_{N-1}) + 2\alpha_N(1 - \alpha_N)(r_N, r_{N-1}). \end{aligned} \quad (15)$$

Differentiating Eq. 15 and solving the resulting equation for  $\alpha_N$  yields the optimum mixing parameter for each iteration as

$$\alpha_N = -\frac{(r_{N-1}, r_N - r_{N-1})}{(r_N - r_{N-1}, r_N - r_{N-1})}. \quad (16)$$

### C. Broyden mixing

Broyden mixing treats the density update as an approximate inversion procedure. The goal is to minimize the difference

$$R[n] = F[n] - n \quad (17)$$

where  $R$  is referred to as the *residual* and  $F$  is defined in Eq. 1. We shall return to the concept of the residual in the next lecture. The exact minimization is, of course, not possible as this would amount to the solution of the original problem. However, an alternative and approximate solution may be devised if we are sufficiently close to the true solution. In order to facilitate this, we make use of the *Jacobian* defined as

$$J \equiv \frac{\delta R}{\delta n} \quad (18)$$

The value of the Jacobian that would result in a zero residual at the  $i + 1$ th step is

$$n_{i+1} = n_i - J^{-1}R_i \quad (19)$$

which we could see easily by writing Eq. 18 as

$$\delta n_i = J_i^{-1}R_i \quad (20)$$

where  $\delta n_i = n_i - n_0$  measures the difference between the density at the current step and the true ground state density. Substituting Eq. 20 into Eq. 19, we have

$$n_{i+1} = n_i - \delta n_i = n_i - (n_i - n_0) = n_0 \quad (21)$$

which means that convergence has been reached.

Of course, the Jacobian can neither be known exactly nor inverted easily. However, the Broyden method proposes a way of making an initial guess for the Jacobian (which is more easily diagonalizable than the original) and iteratively improving it as the iterations progress. To be more precise we don't actually invert at any point during the process the Jacobian itself but we improve the inverse Jacobian progressively instead. At each step  $i$  where

$$\delta n_i = n_i - n_{i-1} = -J_{i-1}^{-1}R_{i-1} \quad (22)$$

the new, improved  $J_i^{-1}$  is chosen such that :

1. The result of the iteration just completed is reproduced, i.e.,

$$0 = \delta n_i - J_i^{-1}(R_i - R_{i-1}). \quad (23)$$

2. The norm of the change in the Jacobian matrix

$$Q = |J_i^{-1} - J_{i-1}^{-1}| \quad (24)$$

is minimized.

The result is

$$J_i^{-1} = J_{i-1}^{-1} \frac{(\delta n_i - J_{i-1}^{-1}R_i)\delta R_i}{\langle \delta R_i | \delta R_i \rangle}, \quad (25)$$

which is suitable for use in small systems but not large systems where the Jacobian from the previous step needs to be stored in its entirety. At this point the *modified Broyden method* comes to rescue where  $Q$  is replaced by a weighted norm

$$Q_{mod} = \sum_{j=1}^i w_j |\delta n_j - J_j^{-1}R_j|^2 + w_0 |J_i^{-1} - J_0^{-1}| \quad (26)$$

### III. HELLMAN-FEYNMAN THEOREM

Once the ground state density has been reached through self-consistent iterations, the Kohn-Sham orbitals corresponding to this density may be used to calculate several properties of the system, one of which is the force on the atoms. This section is somewhat unrelated to the rest of this chapter but it has been placed here for lack of a better home.

Forces on the atoms are a central part of any DFT calculation since these forces, calculated at the end of the relaxation of the electronic coordinates of the are used for Born-Oppenheimer based geometric optimization methods such as static minimization or molecular dynamics.

The Hellman-Feynman theorem is really a general statement about derivatives of expectation values at the ground state but we discuss the specific case of atomic forces. As is known from elementary physics, the force on the  $i$ th atom is given by

$$\vec{F}_i = -\frac{\partial E}{\partial \vec{R}_i} \quad (27)$$

where the electronic energy for the ground state is clearly given by

$$E_e = \langle \Psi | \hat{H} | \Psi \rangle. \quad (28)$$

Substituting Eq. 28 into Eq. 27, we have, for the electronic part of the forces,

$$\vec{F}_{i,e} = -\left\langle \frac{\partial \Psi}{\partial \vec{R}_i} | \hat{H} | \Psi \right\rangle - \langle \Psi | \frac{\partial \hat{H}}{\partial \vec{R}_i} | \Psi \rangle - \langle \Psi | \hat{H} | \frac{\partial \Psi}{\partial \vec{R}_i} \rangle \quad (29)$$

Grouping the first and the third terms together and making use of the relation  $\hat{H}|\Psi\rangle = E|\Psi\rangle$  in Eq. 29 yields

$$\vec{F}_{i,e} = -E \left[ \left\langle \frac{\partial \Psi}{\partial \vec{R}_i} | \Psi \right\rangle + \left\langle \Psi | \frac{\partial \Psi}{\partial \vec{R}_i} \right\rangle \right] - \langle \Psi | \frac{\partial \hat{H}}{\partial \vec{R}_i} | \Psi \rangle = -E \frac{\partial}{\partial \vec{R}_i} \langle \Psi | \Psi \rangle - \langle \Psi | \frac{\partial \hat{H}}{\partial \vec{R}_i} | \Psi \rangle = -\langle \Psi | \frac{\partial \hat{H}}{\partial \vec{R}_i} | \Psi \rangle \quad (30)$$

where in the second term the derivative of the inner product cancels because this inner product yields 1. This is a very important and useful result which states that while differentiating the energy to obtain the forces, we may entirely ignore the change in the wavefunction and just calculate the change in the Hamiltonian operator. In the case of Coulombic DFT, this is further simplified by the fact that the only term that depends on atomic coordinates is the external potential due to atomic nuclei, which luckily is a single-particle operator whose expectation value may be expressed in terms of the single-particle density. As the finishing touch we add to the electronic contribution to the force the ionic contribution to find the total force on the  $i$ th atom. We then finally arrive at

$$\vec{F}_i = - \int d\vec{r} n(\vec{r}) \frac{\partial V_{ext}}{\partial \vec{R}_i} - \frac{\partial E_{II}}{\partial \vec{R}_i} \quad (31)$$

# Lecture XI : Iterative solutions for solving the Kohn-Sham equation

## I. THE DIAGONALIZATION PROBLEM IN PLANEWAVE CALCULATIONS

The Kohn-Sham Hamiltonian when written in the plane-wave basis yields a matrix with a large dimension, which creates a very tight bottleneck in calculation speed since exact diagonalization scales like  $N^3$ . Moreover, in most electronic structure calculations only the first few eigenvalues are needed among the large number of eigenvalues that necessarily result from traditional diagonalization techniques. In order to overcome this difficulty, several alternative diagonalization schemes were either borrowed from linear algebra or devised from scratch by physicists. All of these methods yield approximate solutions to the eigenvalue equation however by careful tuning, results come out to be very close to the exact eigenvalues and eigenvectors. This problem is avoided in localized bases where the size of the basis set is much smaller and direct diagonalization costs little time.

Most of these methods rely on a repeated application of the Hamiltonian operator on an initial guess for an eigenvector and thus creating a sequence of vectors that span a vector space.

## II. ITERATIVE METHODS

The common features of iterative methods are discussed extensively in literature (see, for example, Wood and Zunger *J. Phys. A : Math. Gen.*, **18**, 1985).

The problem we have at hand is finding the  $n$  lowest eigenvalues of a generalized eigenvalue problem

$$\hat{H}|\psi_i\rangle = \lambda \hat{S}|\psi_i\rangle \quad (1)$$

where the dimension of the problem is  $N$  and  $n \ll N$ . In the so-called iterative methods, there are three common basis sets to be determined

- The basis set,  $\{|\phi_i\rangle\}$  that is used in expressing the Hamiltonian,  $\hat{H}$ . In our case, this is the plane-wave basis.
- A complete  $N$ -dimensional set  $\{\psi_i\}$  that spans the entire Hilbert space of the Hamiltonian and by means of which all vectors in the Hilbert space may be expressed.
- A much smaller  $N_b$ -dimensional spanning set  $\{b_i\}$ . The assumption common to all iterative methods is that all eigenvectors of interest may be expanded in terms of this  $N_b$  dimensional vector space.  $N_b$  here satisfies  $n \sim N_b \ll N$ .

In all of the iterative methods, one starts with an  $N_0 \times N_0$  submatrix of the Hamiltonian,  $\hat{H}_0$  and finds its eigenvalues  $\{\lambda_j^0\}$  and eigenvectors  $\{a_j^0\}$  exactly. In the subsequent iterations, these starting eigenvectors are improved iteratively to reach convergence. The approximations to the eigenvalues and eigenvectors are calculated via the solution of the eigenvalue equation

$$\hat{H}^r|c\rangle = \epsilon \hat{S}^r|c\rangle \quad (2)$$

where the reduced operators  $\hat{H}^r$  and  $\hat{S}^r$  are written in the basis  $\{|B_i\rangle\}$  as

$$H_{ij}^r = \langle b_i | H | b_j \rangle \quad (3)$$

$$S_{ij}^r = \langle b_i | S | b_j \rangle \quad (4)$$

(5)

At the  $k$ th step then the approximation to the real eigenvector is given by  $\epsilon_k$  and the approximation to the corresponding eigenvector is given by

$$|a_k\rangle = \sum_i \langle b_i | c_k \rangle |b_i\rangle. \quad (6)$$

After each step new vectors are calculated either to add to the expansion set or directly to improve the approximation to the eigenvector being sought.

Iterative methods may either be *sequential* meaning that a single eigenvector is calculated at a time or in the *block* form where several eigenvalues emerge simultaneously.

Different methods are characterized by their choice of the spanning spaces  $\{|b_i\rangle\}$  and  $\{|\psi_i\rangle\}$ . Some of these methods are briefly discussed below :

### A. The Newton method

The quality of the approximation is measured by a residual vector defined by

$$|R(|A^a\rangle, E^a)\rangle = (\hat{H} - E^a \hat{S})|A^a\rangle. \quad (7)$$

where  $E^a$  and  $|A^a\rangle$  are the approximate eigenvalues and eigenvectors respectively. The residual as defined in Eq. 7 is in turn used to defined the quantity

$$R = \left( \frac{\langle R|R\rangle}{\langle A^a|S|A^a\rangle} \right)^{1/2}. \quad (8)$$

$R$  is a measure of how far the approximation to the eigenvalue and the eigenvector is from the true ones and for the exact solution would be zero. The goal for each iteration then would be to make an addition  $|\delta A\rangle$  to the approximation vector  $|A^a\rangle$  such that the residual is a minimum. This is satisfied by demanding

$$|R(|A^a + \delta A\rangle, E^a)\rangle = |R(|A^a\rangle, E^a)\rangle + (\hat{H} - E^a \hat{S})|\delta A\rangle = 0. \quad (9)$$

The  $|\delta A\rangle$  that satisfies Eq. 9 is given by

$$|\delta A\rangle = -(\hat{H} - E^a \hat{S})^{-1} |R(|A^a\rangle, E^a)\rangle \quad (10)$$

which however takes us back to the initial problem and thus cannot be solved exactly. Instead several methods (such as Newton,...) make a an approximation to the inverse operator in Eq. 10

$$|\delta A\rangle = -D^{-1}|R\rangle \quad (11)$$

where  $D$  is the diagonal of the  $\hat{H} - E^a \hat{S}$ . Thus in the next iteration, the approximate eigenvector is replaced by

$$|A^{new}\rangle = |A^a\rangle + |\delta A\rangle \quad (12)$$

where  $|\delta A\rangle$  is defined in Eq. 11. This update step may be written in basis expansion form as

$$|\delta A\rangle = - \sum_i' \frac{\langle x_i|R\rangle \langle x_i|}{\langle x_i|\hat{H} - E^a \hat{S}|x_i\rangle} \quad (13)$$

where the prime on the sum enforces the omission of such term that give a denominator that is below a certain threshold.

### B. The Lanczos method

The Lanczos method is based on a repeated improvement of the basis set in the following manner. The vectors in the basis set in the  $i$ th step is formed by the repeated application of the Hamiltonian on initial guess vector  $|a_j^0\rangle[1]$ .

$$\{|b_j\rangle^k\} = [|a_j^0\rangle, \hat{H}|a_j^0\rangle, \hat{H}^2|a_j^0\rangle, \dots, \hat{H}^k|a_j^0\rangle] \quad (14)$$

The remaining operations for each iteration are as described above in the Newton method. The vectors in Eq. 14 formed at each iteration must be orthonormalized to all the previous vectors.

### C. The Davidson method

The Davidson method is described by

$$\begin{aligned} \{|x_i\rangle\} &= \{|e_i\rangle\} \\ \{|b_i\rangle^k\} &= [|a_j^0\rangle, |b_j^1\rangle = |\delta A_1\rangle - \sum_{i=1}^{j-1} \langle b_i|\delta A_1\rangle |b_i\rangle] \end{aligned} \quad (15)$$

### III. OPTIMIZATION METHODS

A. Steepest descent

B. Conjugate gradients

---

[1] The iterative vector space thus formed is called a Krylov subspace

## I. THE ELECTRON GAS

We have defined the exchange and correlation energy before as the difference between the exact total energy of a system and the classical Hartree energy. Except in very simple cases it is impossible to treat exchange and correlation exactly and an approximation is in order. The quality of a DFT calculation is determined by how close the approximate exchange and correlation comes to the exact value. In this section, we shall get a glimpse of the -by now- huge literature on different methods of dealing with exchange and correlation in atomistic systems.

Before going into exchange and correlation in real systems, it's important to understand the behavior of a much simpler system, the interacting electron gas not only because it is the system that can be dealt with analytically the most easily but also due to the fact that approximations to exchange and correlation are built upon this simple system.

Let us first start with the noninteracting electron gas in order to to familiarize ourselves with concepts and establish some terminology.

### A. The noninteracting electron gas

A noninteraction electron gas is homogeneous and is completely characterized by its density,  $n = N/V$  where  $N$  is the total number of electrons and  $V$  is the volume of the system. An alternative way of characterizing such a system is through an operational parameter  $r_s$ , which is defined as the radius of a sphere containing a single electron

$$\frac{4\pi}{3}r_s^3 = \frac{1}{N/V} = \frac{1}{n} \Rightarrow r_s = \left(\frac{3}{4\pi n}\right)^{1/3} \quad (1)$$

The density or equivalently  $r_s$  in a real system is of course strongly position-dependent.

The Hamiltonian of the noninteracting electron gas is composed solely of the kinetic energy term and the total energy is just the sum of the kinetic energy of the first  $N$  planewave eigenstates. The *Fermi energy* is defined as the energy of the highest filled planewave state. It is characterized by the *Fermi wavevector* through the simple relation

$$E_F^\sigma = \frac{1}{2}k_F^\sigma \quad (2)$$

where the spin dependence is kept explicitly. The total energy is then just the integral over all the states having wavevectors up to  $\vec{k}_F$ . To simplify the measure, we do not calculate the total kinetic energy but the energy per electron for each spin.

$$\frac{T_{tot}^\sigma}{N} = \frac{\frac{1}{2} \int d\vec{k} k^2}{\int d\vec{k}} = \frac{1}{2} \frac{4\pi \int_0^{k_F^\sigma} k^4}{4\pi \int_0^{k_F^\sigma} k^2} = \frac{1}{2} \frac{(k_F^\sigma)^5/5}{(k_F^\sigma)^3/3} = \frac{3}{5} E_F^\sigma \quad (3)$$

where the last equality is taken from Eq. 2. As seen in the equations above, all the relevant properties of the noninteracting electron system may be determined by its density, which is in turn determined by the total number of electrons.

### B. Interacting electrons

When the interaction between the electrons is turned on, the Hamiltonian becomes

$$\hat{H} = -\frac{1}{2} \sum_i \nabla^2 + \frac{1}{2} \left[ \sum_{i \neq j} \frac{1}{|\vec{r}_i - \vec{r}_j|} - \int d\vec{r} d\vec{r}' \frac{n^2}{|\vec{r} - \vec{r}'|} \right] \quad (4)$$

where the last term has been substituted in place of the potential of the ions simulating the presence of a uniform positive background. It is necessary to include this term in order to cancel the divergence in the electron-electron interaction term. The total energy is then given as

$$E = \langle \hat{H} \rangle = \langle \hat{T} \rangle + \langle \hat{V}_{int} \rangle - \frac{1}{2} \int d\vec{r} d\vec{r}' \frac{n^2}{|\vec{r} - \vec{r}'|} \quad (5)$$

where the difference between the two last terms is precisely the exchange-correlation energy mentioned above.

In the Hartree-Fock method that we have seen before, only the exchange part of the interaction energy is taken into account. Since the problem of the interacting electron gas may be treated analytically at the Hartree-Fock level, we may calculate the exchange energy exactly. We first remind ourselves of the Hartree-Fock equations

$$\epsilon_\nu \phi_\nu(\vec{r}) = \left[ -\frac{1}{2} \nabla^2 + \sum_\lambda \int d\vec{r}' \phi_\lambda^*(\vec{r}') \phi_\lambda(\vec{r}') \frac{1}{|\vec{r} - \vec{r}'|} + V_{\text{ion}} \right] \phi_\nu(\vec{r}) - \sum_\lambda \delta_{\sigma_\lambda \sigma_\nu} \int d\vec{r}' \phi_\lambda^*(\vec{r}') \phi_\nu(\vec{r}') \phi_\lambda(\vec{r}) \frac{1}{|\vec{r} - \vec{r}'|} \quad (6)$$

for the electron system where  $\lambda$  and  $\nu$  are the combined orbital/spin indices. The first term is the usual kinetic energy term, the second is the direct term and the final term is the exchange term. The ionic term, once again, has been substituted with the uniform background charge. Let us first consider the second (direct) term in isolation

$$V_d = \int d\vec{r}' \sum_\lambda |\phi_\lambda|^2 \frac{1}{|\vec{r} - \vec{r}'|} = \int d\vec{r}' \frac{n}{|\vec{r} - \vec{r}'|} \quad (7)$$

There are two things to notice about the result in Eq. 7 : one is that the density is still uniform even though the electrons are allowed to interact and the second is that the direct term is exactly canceled by the potential term due to the uniform background which replaces the ions,

$$V_{bg} = - \sum_I \frac{Z_I}{|\vec{r} - \vec{R}_I|} \longrightarrow -n \int \frac{d\vec{R}}{|\vec{r} - \vec{R}|}. \quad (8)$$

Then we are just left with the kinetic energy and the exchange terms. Surprisingly, the eigenstates of the system of equations in Eq. 6 are also planewaves. The eigenvalues are of course different with the usual free electron kinetic energy being augmented by the exchange term. We shall prove this by substituting directly planewaves in place of the wavefunctions.

$$\epsilon(\vec{k}) \frac{1}{\sqrt{V}} e^{i\vec{k} \cdot \vec{r}} = \frac{k^2}{2} - \frac{1}{V} \frac{1}{\sqrt{V}} \sum_{\vec{k}'} e^{i\vec{k}' \cdot \vec{r}} \underbrace{\int d\vec{r}' \frac{e^{i(\vec{k} - \vec{k}') \cdot \vec{r}'}}{|\vec{r}' - \vec{r}'|}}_I \quad (9)$$

Let's concentrate on the integral labeled  $I$ . This integral is simply the Fourier transform of the Coulombic potential, which is a problem we have dealt with before while we were learning about planewave representation of the Kohn-Sham equations. We shall therefore not go into details but just present the result once again after some rearrangements

$$I \equiv e^{i(\vec{k} - \vec{k}') \cdot \vec{r}} \int d\vec{u} \frac{e^{i(\vec{k} - \vec{k}') \cdot \vec{u}}}{u} = 4\pi \frac{e^{i(\vec{k} - \vec{k}') \cdot \vec{r}}}{|\vec{k} - \vec{k}'|^2}. \quad (10)$$

The last term in Eq. 9 then becomes

$$-\frac{4\pi}{V} \sum_j \frac{e^{i\vec{k}' \cdot \vec{r}} e^{i(\vec{k} - \vec{k}') \cdot \vec{r}}}{|\vec{k} - \vec{k}'|^2} = -\frac{4\pi}{V} \sum_{\vec{k}'} \frac{e^{i\vec{k} \cdot \vec{r}}}{|\vec{k} - \vec{k}'|^2}, \quad (11)$$

which completes the proof that planewaves are really eigenstates. Next we move on to find the eigenvalues. The sum over the wavevectors in Eq. 11 can be converted into an integral in the limit of large total volume. The conversion factor to be used is

$$\sum_j = \frac{V}{(2\pi)^3} \int d\vec{k} \quad (12)$$

Then the sum in Eq. 11 becomes

$$\frac{4\pi}{V} \sum_{\vec{k}'} \frac{1}{|\vec{k} - \vec{k}'|^2} \rightarrow \frac{2}{(2\pi)^2} \int d\phi (-d(\cos \theta)) \int_0^{k_F} \frac{k'^2 dk'}{k^2 - 2kk' \cos \theta + k'^2} \quad (13)$$

with the new choice of the variables for ease of notation. Now let's first evaluate the angular integration.

$$\begin{aligned} \frac{1}{\pi} \int_0^{k_F} \left( -\frac{k'}{2k} \right) dk' \ln \left| k^2 - 2kk' + k'^2 \right|_{-1}^1 &= \frac{1}{2\pi} \int_0^{k_F} dk' \frac{k'}{k} \ln \left| \frac{k^2 + 2kk' + k'^2}{k^2 - 2kk' + k'^2} \right| \\ &= \frac{1}{\pi} \int_0^{k_F} dk' \frac{k'}{k} \ln \left| \frac{k + k'}{k - k'} \right| \end{aligned} \quad (14)$$



The logarithmic integral in Eq. 14 can be evaluated using integration by parts. We shall only present the result here. The  $\vec{k}$ -dependent energy eigenvalue is then

$$\epsilon(\vec{k}) = \frac{k^2}{2} - \frac{k_F}{\pi} f(x) \quad (15)$$

where  $x = \frac{k}{k_F}$  and

$$f(x) = 1 + \frac{1-x^2}{2x} \ln \left| \frac{1+x}{1-x} \right| \quad (16)$$

Note here that the free electron kinetic energy has been augmented by the negative exchange energy. The exchange term then lowers the energy by keeping the electrons of the same spin away from each other thus reducing the Coulomb repulsion.

As we have done before, the total energy of the system in the Hartree-Fock approximation is given by the sum of the single-electron energies corrected by a half of the exchange term. Then

$$\begin{aligned} E_{HF} &= 2 \sum_{\vec{k}} \epsilon(\vec{k}) + \frac{1}{2} \sum_{\vec{k}} \epsilon_{exc}(\vec{k}) \\ &= \frac{2V}{(2\pi)^3} 4\pi \int_0^{k_F} k^2 \left( \frac{k^2}{2} + \frac{1}{2} \epsilon_{exc}(k) \right) \end{aligned} \quad (17)$$

After some rather tedious calculus, we arrive at

$$\frac{E_{HF}}{N} = \left( \frac{2.21}{r_s^2} - \frac{0.916}{r_s} \right) \quad (18)$$

Eq. 18 is an important result which states that for the range of  $r_s$  that are relevant for metals, the exchange term has a very large effect on the total energy so it's an important component needed to describe the behavior of real systems. It has however been pointed out that the Hartree-Fock description has important shortcomings such as the divergence of the effective mass (due to the logarithmic term) and the severe underestimation of cohesive energies in metals. In the case of Li, a more extreme and absurd situation is observed where the Hartree-Fock approximation yields a positive cohesive energy. More surprisingly, it can be proven that the noninteracting picture does much better than Hartree-Fock theory in predicting cohesive energy if the ions are treated more accurately. It is thus evident that the inclusion of exchange by itself deteriorates the already existing result rather than improve it.

The *correlation* energy is defined as the difference between the total exact energy and the sum of kinetic and exchange energies. Correlations are a result of the collective behavior of electrons to screen and decrease the Coulombic interaction. Unlike the exchange term the correlations become more pronounced for opposite spins since they are more likely to occupy nearby locations.

Unfortunately, the correlation contribution cannot be calculated exactly analytically. There have been several propositions for correlation energy terms over the years :

- Wigner : Wigner solved the problem exactly in the limit of low  $r_s$  (the Wigner crystal), assumed the correlation energy would go to a constant in the high  $r_s$  limit and proposed a smooth interpolation inbetween.

$$\epsilon_c = -\frac{0.44}{r_s + 7.8} \quad (19)$$

- Gellman and Breuckner : G. and B. followed a diagrammatic approach where they summed an infinite series of Feynman diagrams and calculated exactly the high  $r_s$  limit.

$$\epsilon_c(r_s) = 0.311 \ln(r_s) - 0.048 + r_s(A \ln(r_s) + C) + \dots \quad (20)$$

The low  $r_s$  is retained and an interpolation is made in between as in Wigner's treatment.

- QMC calculations : Accurate quantum Monte Carlo calculations were made for a series of  $r_s$  values covering the entire relevant range, which was followed by several analytic fits made by different researchers. The most widely used correlation potentials are due the Perdew-Zunger (PZ) and Vosko-Wilkes-Nussair (VWN).

For small  $r_s$  values, which are relevant for metals, the correlation energy is a small contribution. However in the large  $r_s$  limit it becomes very important. So far we have dealt with systems that are unpolarized. For partially polarized systems  $n^\uparrow \neq n^\downarrow$ . For each spin channel, the total exchange energy per electron may be written in terms of the density in a given spin channel as

$$\epsilon_x^\sigma = \frac{E_x^\sigma}{N^\sigma} = -\frac{3}{4\pi} k_F^\sigma = -\frac{3}{4} \left( \frac{6}{\pi} n^\sigma \right)^{1/3} \quad (21)$$

The spin-polarized exchange and correlation energies are usually expressed in terms of the total density and the difference between densities of the two spin channels defined by

$$\xi = \frac{n^\uparrow - n^\downarrow}{n} \quad (22)$$

The total energy of the spin-polarized system is the sum of the exchange energy of the two spin channels. Exchange energy as a function of this difference and the total density is given by

$$\epsilon_x(n, \xi) = \epsilon_x(n, 0) + [\epsilon_x(n, 1) - \epsilon_x(n, 0)] f_x(\xi) \quad (23)$$

where

$$f_x(\xi) = \frac{1}{2} \frac{(1 + \xi)^{4/3} + (1 - \xi)^{4/3} - 2}{2^{1/3} - 1} \quad (24)$$

The polarization dependence of the correlation energy was taken to be the same as that of the exchange energy by PZ while a slightly more accurate treatment was introduced by VWN.

## II. EXCHANGE AND CORRELATION — RULES THAT NEED TO BE OBEYED

### A. The exchange-correlation hole

The fact that exchange and correlation both tend to keep electrons apart gave way to the description of the exchange and correlation contribution in terms of a hole surrounding each electron and keeping other electrons from approaching it. The exchange-correlation hole can be interpreted in relation to the joint probability of finding an electron at point  $\vec{r}$  given that there exists another electrons at point  $\vec{r}'$ . It may be written in terms of a quantity familiar from probabilistic considerations, namely the pair correlation function. The pair-correlation function is determined as the part of the joint probability after the uncorrelated probability has been subtracted away

$$\begin{aligned} n(\vec{r}, \sigma; \vec{r}', \sigma') &= \left\langle \sum_{i \neq j} \delta(\vec{r} - \vec{r}_i) \delta(\sigma - \sigma_i) \delta(\vec{r}' - \vec{r}_j) \delta(\sigma' - \sigma_j) \right\rangle \\ &= N(N-1) \sum_{\sigma_3, \sigma_4, \dots} \int d\vec{r}_3 \dots d\vec{r}_N |\Psi(\vec{r}, \sigma; \vec{r}', \sigma'; \vec{r}_3, \sigma_3; \dots; \vec{r}_N, \sigma_N)|^2 \\ n(\vec{r}, \sigma; \vec{r}', \sigma') &= n(\vec{r}, \sigma) n(\vec{r}', \sigma') + \Delta n(\vec{r}, \sigma; \vec{r}', \sigma') \\ g(\vec{r}, \sigma; \vec{r}', \sigma') &= \frac{n(\vec{r}, \sigma; \vec{r}', \sigma')}{n(\vec{r}, \sigma) n(\vec{r}', \sigma')} = 1 + \frac{\Delta n(\vec{r}, \sigma; \vec{r}', \sigma')}{n(\vec{r}, \sigma) n(\vec{r}', \sigma')} \end{aligned} \quad (25)$$

Defining the density of this hole as  $n_{xc}(\vec{r}, \vec{r}')$ , the exchange-correlation energy may be written using the adiabatic connection formula set forth by Harris (*Physical Review A*, **29**, 1648). The many-particle Hamiltonian may be written in a way to make a connection between noninteracting electrons and the fully interacting case by means of a coupling constant,  $\lambda$ :

$$\hat{H}_\lambda = \hat{T} + \lambda \hat{V}_{ee} + \lambda \hat{V}_{ext}^\lambda \quad (26)$$

where  $\lambda = 0$  corresponds to the noninteracting case and  $\lambda = 1$  to full interaction. The exchange and correlation energy may be written in terms of an average over this coupling constant

$$E_{xc}[n] = \int_0^1 d\lambda \frac{d\langle \Psi_\lambda | V_{ee} | \Psi_\lambda \rangle}{d\lambda} - E_H. \quad (27)$$

In Eq. 27, it is important to keep the density constant as  $\lambda$  is varied. If this is satisfied, all energy terms except the electron-electron interaction term will remain unchanged. By the Hellman-Feynman theorem that we have seen earlier, the expression in Eq. 27 reduces to

$$E_{xc}[n] = \int_0^1 d\lambda \langle \Psi_\lambda | \frac{dV_{ee}}{d\lambda} | \Psi_\lambda \rangle - E_H. \quad (28)$$

The exchange and correlation energy defined in Eq. 28 may also be written in terms of the *average exchange-correlation hole density*,  $\bar{n}_{xc}$  as

$$E_{xc}[n] = \int d\vec{r} d\vec{r}' \frac{n(\vec{r}) \bar{n}_{xc}(\vec{r}, \vec{r}')}{|\vec{r} - \vec{r}'|} \quad (29)$$

## B. Restrictions

Needless to mention, neither the exchange correlation hole density or the pair correlation function may be calculated exactly. The quality of a DFT calculation rests upon how well it approximated these quantities. Many successful approximations have so far been designed making DFT a practical and powerful tool to predict several properties of many systems. While devising such approximations there are many restrictions that need to be obeyed. These all stem from the definition of the exact exchange and correlation hole.

1. **Sum rule** : The exchange and correlation hole as defined in Eq. 25 integrates over space exactly to -1 per electron.

$$\begin{aligned} \sum_{\sigma\sigma'} \int d\vec{r} d\vec{r}' n_{xc}(\vec{r}, \vec{r}') &= N(N-1) \int d\vec{r} d\vec{r}' d\vec{r}_3 \cdots d\vec{r}_N |\psi(\vec{r}, \sigma, \vec{r}', \sigma', \vec{r}_3, \sigma_3 \cdots, \vec{r}_N \sigma_N)|^2 \\ &\quad - N^2 \sum_{\sigma, \sigma_2, \cdots, \sigma_N} \int d\vec{r} d\vec{r}_2 \cdots d\vec{r}_N |\psi(\vec{r}, \sigma, \vec{r}_2, \sigma_2 \cdots, \vec{r}_N \sigma_N)|^2 \\ &\quad \times \sum_{\sigma', \sigma_2, \cdots, \sigma_N} \int d\vec{r}' d\vec{r}_2 \cdots d\vec{r}_N |\psi(\vec{r}', \sigma, \vec{r}_2, \sigma_2 \cdots, \vec{r}_N \sigma_N)|^2 \\ &= N^2 - N - N^2 = -N \end{aligned} \quad (30)$$

The exchange-correlation hole may be split into separate exchange and correlation contributions.  $n_x$  is the part of the hole that describes the tendency of like spins to avoid each other. The exchange hole must also integrate to -1, a fact which comes from the fact that an electron cannot be at  $\vec{r}'$  if it is at  $\vec{r}$ . Because of this and Eq. 30 the remaining part of  $n_{xc}$ , a.k.a. the correlation hole must integrate to zero.

2. **Scaling properties** : Scaling properties (as many other exact constraints) for the exact functionals were obtained by Levy and Perdew (*Physical Review A*, **32**, 2010). We give them here without proof :

$$E_x[n_\lambda] = \lambda E_x[n] \quad (31)$$

$$E_c[n_\lambda] > \lambda E_c[n] \quad \text{for } \lambda > 1 \quad (32)$$

$$E_c[n_\lambda] < \lambda E_c[n] \quad \text{for } \lambda < 1 \quad (33)$$

3. **The one-electron limit** : In the one-electron limit, the exchange energy still exists due to the fact that the sum in the exchange term in any Slater determinant formalism is designed to contain the term  $i = j$ . Moreover, this term is exactly equal to the negative Hartree term and they cancel each other exactly. This just leaves the correlation term which should be equal to zero if we want to eliminate any spurious self-interaction terms. Thus if  $n^{(1)}$  is the one-electron density,

$$E_c[n^{(1)}] = 0 \quad \text{and} \quad (34)$$

$$E_x[n^{(1)}] = -E_H[n^{(1)}] \quad (35)$$

$$(36)$$

4. **The Lieb-Oxford bound** : The exact exchange-correlation energy satisfies

$$E_{xc}[n] \geq -1.68 \int d\vec{r} n(\vec{r})^{4/3} \quad (37)$$

5. **The derivative discontinuity** : This property is very important in describing the gap in insulators. This so-called *fundamental gap* is defined to be the difference in energy of the same system with number of electrons differing by one :

$$E_g = (E_{N+1} - E_N) - (E_N - E_{N-1}) \quad (38)$$

This together with Koopman's theorem says that the gap is equal to the difference between the HOMO (highest-occupied molecular orbital) and LUMO (lowest unoccupied molecular orbital) obtained from a Kohn-Sham calculation. In terms of the eigenvalues then the true gap should be given as

$$E_g = \epsilon_{N+1}(N+1) - \epsilon_N(N) \quad (39)$$

where what's usually available to us is the difference in HOMO and LUMO that we obtain in an  $N$ -electron calculation

$$\epsilon_g = \epsilon_{N+1}(N) - \epsilon_N(N). \quad (40)$$

We shall call the difference  $\Delta_{xc} = E_g - \epsilon_g = \epsilon_{N+1}(N+1) - \epsilon_{N+1}(N)$  anticipating that the difference should arise from exchange and correlation.

The variational principle applied together with the constraint that the density integrates to the total number of electrons dictates that

$$\frac{\partial T_{s,M}}{\partial n} + V = \mu_M \quad (41)$$

where  $M$  is the total number of electrons. The chemical potential  $\mu_M$  is the energy of the eigenenergy of the  $M$ th orbital for all  $M$  except  $N$ . Taking the difference of the two limits where  $M$  tends to  $N$  from above and to  $N+1$  from below

$$\epsilon_g = \frac{\partial T_s}{\partial \delta n_+} - \frac{\partial T_s}{\partial \delta n_-} \quad (42)$$

Thus there is a discontinuity in the single-particle kinetic energy. A similar discontinuity exists in the exchange-correlation potential with

$$\Delta_{xc} = V_{xc}^+ - V_{xc}^- \quad (43)$$

where

$$V_{xc}^\mp = \frac{\partial E_{xc}}{\partial n_\mp} \quad (44)$$

This conclusion explain why *local density approximation* to exchange and correlation, the gap is severely underestimated. In LDA (besides the question of an accurate  $\epsilon_g$ ),  $\Delta_{xc}$  is identically zero.

### III. APPROXIMATIONS TO EXCHANGE AND CORRELATION

#### A. Local density approximation(LDA)

The oldest and the most popular of the exchange-correlation functional, LDA assumes a simple form which is a linear functional of the density

$$E_{xc}^{LDA} = \int d\vec{r} n(\vec{r}) \epsilon_{xc}^{hom}(n(\vec{r})) \quad (45)$$

where  $\epsilon_{xc}^{hom}(n(\vec{r}))$  is the exchange correlation energy density of an interacting homogeneous electron gas at the density  $n(\vec{r})$ . Recall here that the exchange part of  $\epsilon_{xc}^{hom}(n(\vec{r}))$  is calculated exactly in the Hartree-Fock sense whereas the correlation part is a result of a fit to accurate MC data. Although LDA is a very crude approximation for systems that are not as homogeneous as an electron gas, it has proven to be a huge success much to the surprise of the community. This is in part due to the cancellation of errors where LDA typically overestimates  $E_X$  while it underestimates  $E_c$ . This comes from the fact that while the LDA xc hole satisfies the sum rule described above, the individual exchange and correlation holes do

not. This can only occur if the integrated errors in the exchange hole cancels that in the correlation hole. LDA has a big drawback which is that it does not satisfy the one-electron limit. This may be fixed by augmenting it with what is called a *self-interaction correction (SIC)*. The correction

$$E_{xc}^{LDA,SIC} = E_{xc}^{LDA} - \sum_i (E_H[n_i] - E_{xc}[n_i]) \quad (46)$$

subtracts the error which would be made by the Hartree and xc functionals in the case of a single electron.

## B. Generalized gradient approximations

LDA treats all systems as homogeneous. However, real systems are clearly inhomogeneous with spatially varying electric fields due to nuclei and screening. In order to take this into account, a step may be taken beyond the LDA by including the derivative information of the density into the exchange-correlation functionals. One way to do this is to try to include gradients  $|\nabla n(\vec{r})|$ ,  $|\nabla n(\vec{r})|^2$ ,  $\nabla^2 n(\vec{r})$   $\dots$  into the approximation. However, what works better is to seek a more general functional

$$E_{xc}^{GGA} = \int d\vec{r} f(n(\vec{r}), \nabla n(\vec{r})). \quad (47)$$

Such functionals are referred to as generalized gradient approximations. There are different flavors of GGA, each making a different choice for the function  $f$ .

## Homework I : Variational Principle and Concepts in Solid State

### Question I (40 pnts.) : The variational solution for the double well

Consider the double-well potential given by

$$V = (x^2 - 1)^2$$

which is plotted in the graph to the right.

a. Using the basis

$$\chi_1 = e^{-(x-1)^2}$$

$$\chi_2 = e^{-(x+1)^2}$$

which includes the superposition of two possible wavefunction forms that are centered on either of the two wells,

- form the  $2 \times 2$  Hamiltonian and overlap operators. **(10 pnts.)**
- diagonalize the resulting generalized eigenvalue equation as discussed in class and find the ground state energy. **(5 pnts.)**
- plot the ground state and the other excited state wavefunctions using some sort of software and discuss these wavefunctions in the context of *bonding* and *anti-bonding* states you will remember from chemistry. **(5 pnts.)**

b. Enlarge your basis by including the wavefunctions

$$\chi_3 = (x - 1)e^{-(x-1)^2}$$

$$\chi_4 = (x + 1)e^{-(x+1)^2}$$

Using this new basis

- download the codes entitled `integrate.m` and `differentiate.m` from the web site. Using `Octave` or `Matlab` and these codes, form the  $4 \times 4$  Hamiltonian and overlap matrices. **(10 pnts.)**

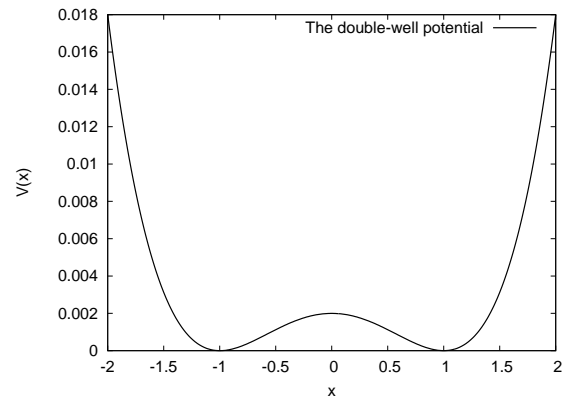
Example usage from `Octave` :

```
octave:1> x=-5:0.01:5;
octave:2> chi1=exp( (x-1).^2 );
octave:3> chi1p=differentiate(x,chi1);
octave:4> T(1,1)=integrate(x,0.5*chi1p.*chi1p);
octave:5> V(1,1)=integrate(x,(x.^2-1).^2*chi1p.*chi1p);
octave:6> S(1,1)=integrate(x,chi1p.*chi1p);
      .
      .
      .
octave:xx> [aa,bb,q,z,v,w,lambda]=qz(T+V,S);
```

Here the last line solves the generalized eigenvalue equation. `v` and `w` contain the eigenvectors (they should be equal when all the matrices involved are Hermitian, which is the case in our problem) and `lambda` contains the eigenvalues.

- find the ground state energy and compare with the energy from the previous part. Is the ground state energy as expected? What can you say the symmetry of the wavefunction that is allowed by this new, extended basis? **(10 pnts.)**

NOTE that you are of course free to use any software you want or solve it analytically. If you are going to do it analytically, I suggest that you make use of the parity (even/odd) of the wavefunctions to reduce some matrix elements to zero.



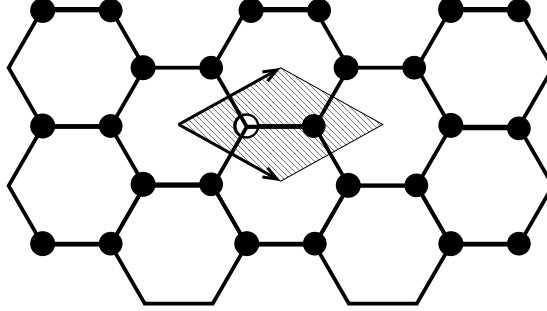
c. **(Bonus : 10 pnts.)** Write a code to generalize this problem to a basis set that has the expression

$$\chi_n = (x-1)^{(n-1)/2} e^{-(x-1)^2}$$

$$\chi_{n+1} = (x+1)^{(n-1)/2} e^{-(x-1)^2}$$

where  $n = 1, \dots, N$  for an arbitrary  $N$ . Describe how the ground state and the wavefunctions evolve.

**Question II (30 pnts.) : The BZ of graphene**



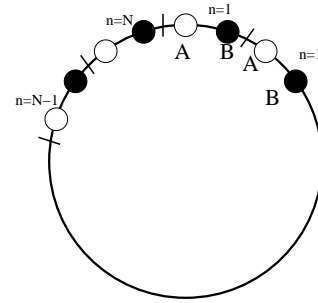
- Looking at the representation of a two-dimensional graphene sheet above, write down the primitive unit cell lattice vectors and basis vectors. **(5 pnts.)**
- Show (with a sketch) that the first BZ of graphene is a hexagon which is rotated with respect to the real space arrangement of hexagons. **(10 pnts.)**
- Now find a larger unit cell in real space. Show (again with a sketch) that the first BZ of this new unit cell is included in the first BZ of the primitive unit cell. **(15 pnts.)**

**Question III (30 pnts.): Bloch's theorem**

Assume that we have a very simple system made of  $N$  units of two different kinds of atoms,  $A$  and  $B$ , that are arranged around a ring as seen in the figure. In this case, obviously, there are two atoms in the unit cell and the system obeys Born-von Karman boundary conditions. Assume that there is only nearest neighbor interaction.

Now, let's form a variational problem, proposing two kinds of wavefunctions : a single state centered on each atom residing on site  $A$ ,  $\chi_{A,n}$  and a single state for atoms residing on site  $B$ ,  $\chi_{B,n}$  where  $n$  ranges from 1 to  $N$ .

As discussed in class, there are two ways of solving this problem : one is to write a Hamiltonian and overlap operators with the elements



$$\begin{aligned}\langle \chi_{A,n} | \chi_{A,m} \rangle &= \langle \chi_{B,n} | \chi_{B,m} \rangle = \delta_{nm} \\ \langle \chi_{A,n} | \chi_{B,m} \rangle &= 0 \\ \langle \chi_{A,n} | H | \chi_{A,m} \rangle &= \varepsilon_A \delta_{nm} \\ \langle \chi_{B,n} | H | \chi_{B,m} \rangle &= \varepsilon_B \delta_{nm} \\ \langle \chi_{A,n} | H | \chi_{B,m} \rangle &= \begin{cases} t & \text{if } A \text{ and } B \text{ are nearest neighbors} \\ 0 & \text{otherwise} \end{cases}\end{aligned}$$

and the other is to make use of Bloch's theorem to reduce the dimensionality of the problem.

- For the set of localized functions above, write down (no need to diagonalize) the Hamiltonian matrix for  $N = 5$ . **(5 pnts.)** One can imagine diagonalizing the Hamiltonian for any given  $N$  (this time we just have a regular diagonalization problem since the overlap matrix is identity) and finding the eigenfunctions.

- b. Show that a general, *delocalized*,  $k$ -dependent wavefunction that has the form

$$\chi_A^k = \frac{1}{\sqrt{N}} \sum_{n=1}^N e^{ink} \chi_{A,n}$$

$$\chi_B^k = \frac{1}{\sqrt{N}} \sum_{n=1}^N e^{ink} \chi_{B,n}$$

satisfies Bloch's theorem. **(10 pts.)**

- c. Using this new basis, the Hamiltonian reduces to a  $2 \times 2$  matrix, but which now must be solved at every  $k$ . Write down this  $2 \times 2$  matrix and solve for a general  $k$ . Draw the band structure, which has two branches. **(15 pts.)**



## Homework II : Hartree-Fock theorem

### Question I (20 pnts.) : Diagonalizing the Lagrange multipliers

The Hartree-Fock equations that emerge as a result of minimizing the energy with respect to the orbitals are

$$\hat{h}_{\text{eff}}|\phi_i\rangle = \sum_{ij} \Lambda_{ij}|\phi_j\rangle$$

where  $\Lambda$  is a matrix whose elements are the Lagrange multipliers that keep the orbitals orthonormal during the minimization process.

- (a) Use this equation to show that  $\Lambda$  is Hermitian.
- (b) If  $\Lambda$  is Hermitian there is a unitary transformation  $U$  under which it is diagonal. Using this transformation, define a new set of orbitals

$$|\phi'_i\rangle = \sum_{j=1}^N U_{ij}|\phi_j\rangle$$

and show that

$$|\phi'_i\rangle = \sum_{j=1}^N U_{ij}|\phi_j\rangle.$$

NOTE that you may choose to follow a different path while solving this question.

### Question II (40 pnts.) : Spin properties of Slater determinants

Consider a two-electron system with the explicitly spin-dependent Slater determinant

$$\Phi = \frac{1}{2}[\phi^+(\vec{r}_1)\alpha(1)\phi^-(\vec{r}_2)\beta(2) - \phi^+(\vec{r}_2)\alpha(2)\phi^-(\vec{r}_1)\beta(1)]$$

where  $\alpha$  and  $\beta$  are the spin part of the orbitals and  $\phi^\mp$  depend only on the spatial coordinate  $\vec{r}$ . In the general case  $\phi^+$  and  $\phi^-$  are allowed to be different. Prove that for the general case  $\Phi$  is always an eigenfunction of  $\hat{S}_z$  but an eigenfunction of  $\hat{S}^2$  only if  $\phi^+ = \phi^-$ .

#### Hints :

- a. The spin operators are given by

$$\hat{S}_x = \sum_{i=1}^N \hat{s}_{i,x} \quad \hat{S}_y = \sum_{i=1}^N \hat{s}_{i,y} \quad \hat{S}_z = \sum_{i=1}^N \hat{s}_{i,z} \quad (1)$$

and

$$\hat{S}^2 = \hat{S}_x^2 + \hat{S}_y^2 + \hat{S}_z^2$$

- b. The application of  $\hat{s}_z$  on the spin parts of the operators yield

$$\begin{aligned} \hat{s}_z\beta &= -\beta \\ \hat{s}_z\alpha &= \alpha \end{aligned}$$

- c. You may make use of the ladder operators.

**Question III (40 pnts.) : The Thomas-Fermi approximation**

A predecessor of the modern DFT is the so-called Thomas-Fermi approximation where the kinetic energy is also written in terms of the density, where locally the kinetic energy is considered to be the same as that of a homogeneous electron gas

$$T[n] = \left(\frac{3}{10}\right) (3\pi^2)^{2/3} \int d\vec{r} n^{5/3}(\vec{r}) \quad (2)$$

and the Coulombic part is taken only to be the Hartree energy

$$E_H[n] = \frac{1}{2} \int \int \frac{n(\vec{r})n(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r} d\vec{r}'.$$

- a. For an arbitrary external potential  $v_{ext}(\vec{r})$ , derive the Thomas-Fermi equation by minimizing the total energy with respect to the density under the constraint that the density integrates to the number of atoms.
- b. Show that the (single) Lagrange multiplier is the chemical potential,  $\mu$ .

## Homework III-IV : N-representability, planewave expansion, Kohn-Sham orbitals, pseudopotentials

### Question I (50 pnts.) : Janak's theorem

Read Janak's paper (Phys Rev B **18** 7165 (1978)).

- (25 pnts) Reproduce the proof outlined between Eqs. 6 and 10. Fill in the inbetween steps and explain your work.
- (25 pnts) Prove the result in Eq. 12, once again explaining your work well.

**Question III (50 pnts.) Hartree and XC potential in planewaves** In lecture, we showed how to evaluate terms in the total energy in the Fourier representation and talked about the advantages of such a representation. In order to get to calculating the energy in planewave representation, however, we need to first complete writing the Kohn-Sham Hamiltonian in terms of planewaves. Following the derivation in lecture, do the following

- (30 pnts.) Show that the Hartree potential in the Fourier representation is

$$V_H(\vec{G}) = 4\pi \frac{n(\vec{G})}{G^2}$$

- (20 pnts.) Show that the exchange-correlation potential in the Fourier representation is

$$V_{xc}(\vec{G}) = \sum_{\vec{G}'} n_{xc}(\vec{G} - \vec{G}') \epsilon'_{xc}(\vec{G}') + \epsilon_{xc}(\vec{G})$$

The meaning of the terms should be obvious. If they are not, consult previous notes.

### Question III (50 pnts.) : Oscillatory and sharp functions

Pseudopotentials ensure two things :

- Cut down the number of electrons in the calculation.
- Make the external potential *smoother* and the pseudowavefunctions *less oscillatory near the origin*.

The second point is necessary to reduce the number of planewaves necessary in the planewave expansion. In this exercise we will illustrate the size of the Fourier components necessary for a sharp vs smooth and an oscillatory versus smooth function.

- (10 pnts.) Consider two functions in one dimension, one oscillatory and one smooth :

$$\begin{aligned} f(x) &= C_1 e^{-x^2} \\ g(x) &= C_2 e^{-x^2} \cos(qx) \end{aligned}$$

where  $C_1$  and  $C_2$  are normalization constants and  $q$  is an arbitrary parameter.

- (5 pnts.) Normalize both functions so that  $\int dx f^2(x) = \int dx g^2(x) = 1$ . Then plot them *on the same graph* to verify that on the same scale one is very oscillatory while one is very smooth.
- (15 pnts.) Take the Fourier transform of both functions and obtain the Fourier components  $\tilde{f}(k)$  and  $\tilde{g}(k)$
- (10 pnts.) Now plot  $\tilde{f}(k)$  and  $\tilde{g}(k)$  versus  $k$  to verify that for  $q > 1$  the non-zero Fourier components of the oscillatory wavefunction has a larger domain than the smooth function. Moreover, show that this domain increases as  $q$  increases.

Hints :

- Do the integrations first for  $g(x)$  for an arbitrary  $q$ , then to deal with  $f(x)$  just take  $q$  to zero.
- Your integrations will always be in the form of a completion of exponents to squares followed by Gaussian integrations.

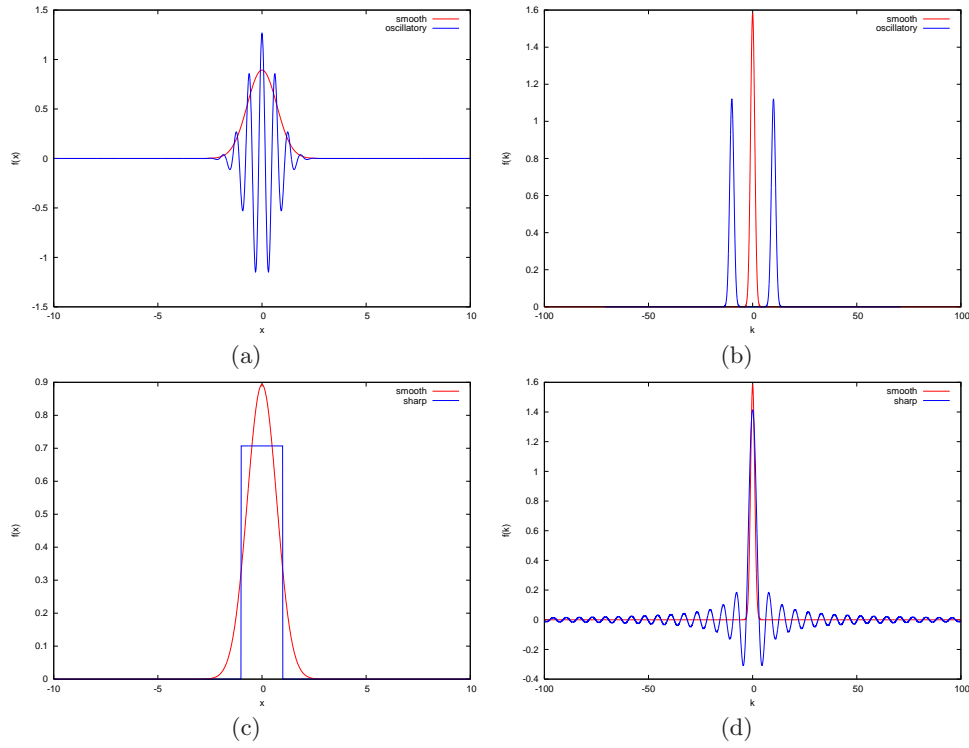


FIG. 1: The smooth and oscillatory(sharp) function in real (a,c) and Fourier (b,d) space

b. (10 pnts.) Consider once again the Gaussian from the previous part together with a very sharp step function :

$$f(x) = C_1 e^{-x^2}$$

$$g(x) = C_2 \begin{cases} 1 & \text{if } |x| < L \\ 0 & \text{otherwise} \end{cases}$$

where  $2L$  is the width of the step function. Repeat the steps above for this case (normalize, plot, take Fourier transform, plot) to see that the nonzero Fourier components of the sharp function has a much larger range than those of the smooth function. Obviously you do not have to repeat the calculations for the Gaussian.

As a result, you should expect to get something like the plots in Fig. 1.

**Question IV (50 pnts.) : Harmonic atom (Adapted from Kaxiras Chapter 2, problem 10)**

Consider a fictitious atom with a harmonic radial potential

$$\left[ -\frac{\hbar^2}{2m_e} \frac{d^2}{dr^2} + \frac{1}{2} m_e \omega^2 r^2 \right] \phi(r) = \epsilon_i \phi(r)$$

a. (20 pnts.) Write a pseudowavefunction for the second state

$$\phi_2(r) = \sqrt{\frac{2\alpha}{\sqrt{\pi}}} \left( \alpha^2 r^2 - \frac{1}{2} \right) e^{-\alpha^2 r^2 / 2}$$

assuming the following form

$$\phi_2^{PS}(r) = \begin{cases} A z^2 e^{-B z^2} & r \leq r_c \\ \phi_2(r) & r > r_c \end{cases}$$

where  $z = r \sqrt{m_e \omega / \hbar}$ , determine the parameters  $A$  and  $B$  such that the pseudowavefunction and its derivative is continuous at  $r_c$ .

- b. **(20 pnts.)** Invert the Schrödinger equation for this pseudowavefunction and find the corresponding pseudopotential.
- c. **(10 pnts.)** Plot the pseudowavefunction and the pseudopotential against their counterparts, taking  $r_c$  to coincide with the outermost bump in the real wavefunction.

Hint : Do NOT attempt to solve a Kohn-Sham equation. This problem is just a noninteracting particle problem.

## Homework IV : Pseudopotential and DFT codes

Please send the Octave functions to me or Sibel by EMAIL.

### Question I (50 pnts) : Pseudopotential code

In class, we constructed a pseudowavefunction for the 4th state of the harmonic oscillator.

- (a) Write a second set of codes to do the same for the 2nd state. In order to complete this assignment, you will need to come up with a smooth function to fit your 2nd state to and also do some math to figure out the fitting parameters. Hand in this math (on paper) together with your codes.
- (b) Invert the one-particle Schrödinger equation to find a pseudopotential for the 2nd and the 4th states. Plot these pseudopotentials.

### Question II (50 pnts) : The DFT code

In class, we wrote a complete one-dimensional DFT code.

- (a) Write a function `get_total_energy.m` with the following prototype to give you the total energy at each step of the SCF cycle.

```
E=get_total_energy(n,VH,Vxc,...)
```

You may or may not need more arguments depending on how you construct your function. Once you have written this function, add it to your `dft.m`. Hand in both the energy code and the modified `dft.m`.

- (b) Write a function `get_VH_2D.m` that finds the Hartree energy not in one-dimension but in two-dimensions. You DO NOT need to integrate this into the rest of the code, just a stand-alone function should do.