

Lecture II

Contents

I. The many-body Hamiltonian	1
A. Atomic units	3
II. The functional derivative	3
III. The definition of density	5
IV. Energy in terms of the density	5
V. The Hohenberg-Kohn theorem	8
A. Some important issues related to Hohenberg-Kohn theorems	9
1. Degenerate ground states	9
2. N- and V-representability, Lieb-Levy constrained search formalism	9
3. Fractional Total Particle Number, Discontinuous Derivatives	12
VI. Kohn-Sham equations	13
A. Issues regarding the Kohn-Sham methods	14
1. Meaning of Kohn-Sham eigenvalues	14
2. Fractional Occupations and Janak's Theorem	14
References	15

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,

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

where m_e is the electron mass, ϵ_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

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

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*.

A. 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 ϵ_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

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

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

II. 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 :

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

$$F[f(x)] : f(x) \rightarrow y \quad \text{functional} \quad (11)$$

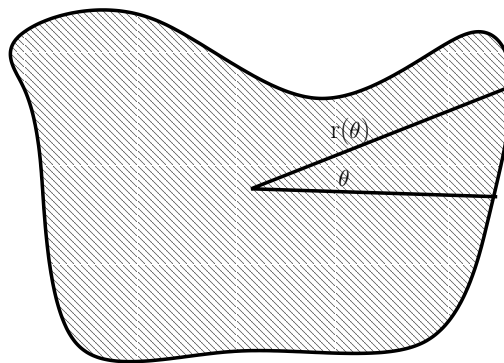
As seen in Eq. 11, arguments of functionals are written in square brackets to distinguish them from regular functions. The perimeter and area of a closed curve parametrized by an angle are familiar examples of functionals.

$$\begin{aligned} P[r] &= \int_0^{2\pi} r(\theta) d\theta \\ A[r] &= \frac{1}{2} \int_0^{2\pi} r^2(\theta) d\theta \end{aligned} \quad (12)$$

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

$$F[f(x) + \delta f(x)] - F[f(x)] = \int dx \frac{\delta F[f(x)]}{\delta f(x)} \delta f(x) \quad (13)$$

Taking for example the area of a parametrized curve defined in Eq. 12, we evaluate its functional derivative with respect to r using the definition in Eq. 13



$$\begin{aligned} A[r + \delta r] - A[r] &= \int \frac{1}{2} ((r + \delta r)^2 - r^2) d\theta = \int \frac{1}{2} (r^2 + 2r\delta r - r^2) d\theta \\ &= \int r \delta r d\theta \Rightarrow \frac{\delta A[r]}{\delta r} = r \end{aligned} \quad (14)$$

Functionals that are written as an integral of a simple function of their argument (such as $P(\theta)$ and $A(\theta)$ in Eq. 12) 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

$$F[f(x)] = \int_a^b dx a(f(x)) \quad (15)$$

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

$$\int dx [a(f(x) + \delta f(x)) - a(f(x))] = \int dx \left[a(f) + \frac{da}{df} \delta a(x) - a(f) \right] = \int dx \frac{da}{df} \delta a(x) \Rightarrow \frac{\delta a}{\delta f} = a'(f(x)) \quad (16)$$

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

$$\begin{aligned}
F[f, f'] &= \int dx [a(f(x) + \delta f(x), f'(x) + \delta f'(x)) - a(f(x), f'(x))] \\
&= \int dx \left[a(f(x), f'(x)) + \frac{\partial a}{\partial f} \delta f(x) + \frac{\partial a}{\partial f'} \delta f'(x) - a(f(x), f'(x)) \right] \\
&= \int dx \frac{\partial a}{\partial f} \delta f(x) + \int dx \frac{d}{dx} \frac{\partial a}{\partial f'} \delta f(x) - \int dx \frac{d}{dx} \frac{\partial a}{\partial f'} \delta f(x)
\end{aligned} \tag{17}$$

In the last line of Eq. 17, the second term cancels due to the assumption of vanishing derivatives. The Second and third terms combine to give the functional derivative as

$$\frac{\delta F[f, f']}{\delta f} = \frac{\partial a}{\partial f} - \frac{d}{dx} \frac{\partial a}{\partial f'} \tag{18}$$

It is straightforward to generalize this formula to a nonlocal functional

$$F[f, f', f'', f''', \dots] = \int a(f, f', f'', f''', \dots) dx \tag{19}$$

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

$$\frac{\delta F[f, f', f'', f''', \dots]}{\delta f} = a - \frac{d}{dx} \frac{\partial a}{\partial f'} + \frac{d^2}{dx^2} \frac{\partial a}{\partial f''} - \frac{d^3}{dx^3} \frac{\partial a}{\partial f'''} + \dots \tag{20}$$

III. 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

$$\hat{n}(\vec{r}) = \sum_{i=1, N} \delta(\vec{r} - \vec{r}_i) \tag{21}$$

$$\begin{aligned}
n(\vec{r}) &= \langle \Psi | \hat{n}(\vec{r}) | \Psi \rangle = \sum_{i=1, N} \int \delta(\vec{r} - \vec{r}_i) |\Psi(\vec{r}_1, \dots, \vec{r}_N)|^2 d\vec{r}_1 \dots d\vec{r}_N \\
&= \int |\Psi(\vec{r}, \vec{r}_2, \dots, \vec{r}_N)|^2 d\vec{r}_2 d\vec{r}_3 \dots d\vec{r}_N + \int |\Psi(\vec{r}_1, \vec{r}, \dots, \vec{r}_N)|^2 d\vec{r}_1 d\vec{r}_3 \dots d\vec{r}_N + \dots \\
&= N \int |\Psi(\vec{r}, \dots, \vec{r}_N)|^2 d\vec{r}_2 \dots d\vec{r}_N
\end{aligned} \tag{22}$$

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

$$\int d\vec{r} n(\vec{r}) = N. \tag{23}$$

IV. 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

$$\begin{aligned}\hat{H}_e &= \hat{T} + \hat{V}_{en} + \hat{V}_{ee} \\ &= -\frac{1}{2} \sum_i^{N_e} \nabla_i^2 - \sum_i^{N_e} \sum_I^{N_n} \frac{Z_I}{|\vec{r}_i - \vec{R}_I|} + \frac{1}{2} \sum_i^{N_e} \sum_{j \neq i}^{N_e} \frac{1}{|\vec{r}_i - \vec{r}_j|}\end{aligned}\quad (24)$$

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

$$\langle \Psi(\vec{r}_1, \dots, \vec{r}_n) | \hat{V}_{ne} | \Psi(\vec{r}_1, \dots, \vec{r}_n) \rangle = - \sum_i^{N_e} \sum_I^{N_n} \int \Psi^*(\vec{r}_1, \dots, \vec{r}_N) \frac{Z_I}{|\vec{r}_i - \vec{R}_I|} \Psi(\vec{r}_1, \dots, \vec{r}_N) d\vec{r}_1 \dots d\vec{r}_N \quad (25)$$

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

$$\langle \Psi(\vec{r}_1, \dots, \vec{r}_n) | \hat{V}_{ne} | \Psi(\vec{r}_1, \dots, \vec{r}_n) \rangle = - \sum_i^{N_e} \sum_I^{N_n} \int \frac{Z_I}{|\vec{r}_i - \vec{R}_I|} |\Psi(\vec{r}_1, \dots, \vec{r}_N)|^2 d\vec{r}_1 \dots d\vec{r}_N \quad (26)$$

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

$$\begin{aligned}\langle \Psi(\vec{r}_1, \dots, \vec{r}_n) | \hat{V}_{ne} | \Psi(\vec{r}_1, \dots, \vec{r}_n) \rangle &= - \sum_I^{N_n} \left[\int \frac{Z_I}{|\vec{r}_1 - \vec{R}_I|} |\Psi(\vec{r}_1, \dots, \vec{r}_N)|^2 d\vec{r}_1 \dots d\vec{r}_N \right. \\ &\quad \left. + \int \frac{Z_I}{|\vec{r}_2 - \vec{R}_I|} |\Psi(\vec{r}_1, \dots, \vec{r}_N)|^2 d\vec{r}_1 \dots d\vec{r}_N + \dots \right]\end{aligned}\quad (27)$$

For each of the N_e terms in the sum in Eq. 27, we separate the integral over the variable in the Coulombic terms from the others.

$$\begin{aligned}\langle \Psi(\vec{r}_1, \dots, \vec{r}_n) | \hat{V}_{ne} | \Psi(\vec{r}_1, \dots, \vec{r}_n) \rangle &= - \sum_I^{N_n} \left[\int \frac{Z_I}{|\vec{r}_1 - \vec{R}_I|} d\vec{r}_1 \int |\Psi(\vec{r}_1, \dots, \vec{r}_N)|^2 d\vec{r}_2 d\vec{r}_3 \dots d\vec{r}_N \right. \\ &\quad \left. + \int \frac{Z_I}{|\vec{r}_2 - \vec{R}_I|} d\vec{r}_2 \int |\Psi(\vec{r}_1, \dots, \vec{r}_N)|^2 d\vec{r}_1 d\vec{r}_3 \dots d\vec{r}_N + \dots \right]\end{aligned}\quad (28)$$

For each term in Eq. 28, the second integral is the definition of the density, $n(\vec{r})$ as given in Eq. 22, that is to say Eq. 28 may be written as

$$E_{ne} = \langle \Psi(\vec{r}_1, \dots, \vec{r}_n) | \hat{V}_{ne} | \Psi(\vec{r}_1, \dots, \vec{r}_n) \rangle = -\frac{1}{N_e} \sum_I^{N_n} \left[\int \frac{Z_I}{|\vec{r}_1 - \vec{R}_I|} n(\vec{r}_1) + \int \frac{Z_I}{|\vec{r}_2 - \vec{R}_I|} n(\vec{r}_2) + \dots \right] \quad (29)$$

Since for each term in Eq. 29, an integral over the argument of the density is taken, we may replace these variables by a dummy variabel, say $v\vec{r}$. 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

$$E_{ne} = - \sum_I^{N_n} \int n(\vec{r}) \frac{Z_I}{|\vec{r} - \vec{R}_I|} d\vec{r} = \int n(\vec{r}) V_{ne}(\vec{r}) d\vec{r} \quad (30)$$

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

$$E_{ee} = \frac{1}{2} \int \int d\vec{r} d\vec{r}' \frac{n^{(2)}(\vec{r}, \vec{r}')}{|\vec{r} - \vec{r}'|} \quad (31)$$

where $n^{(2)}$ is given by

$$n^{(2)}(\vec{r}, \vec{r}') = \frac{N(N-1)}{2} \int d\vec{r}_3 d\vec{r}_4 \cdots d\vec{r}_N |\Psi(\vec{r}, \vec{r}', \dots, \vec{r}_N)|^2 \quad (32)$$

and 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. 31 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. 31 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

$$n^{(2)}(\vec{r}, \vec{r}') = n(\vec{r})n(\vec{r}') + \Delta n^{(2)}(\vec{r}, \vec{r}') \quad (33)$$

The electron-electron energy of Eq. 31 may then be written as

$$E_{ee} = \frac{1}{2} \int \int d\vec{r} d\vec{r}' \frac{n(\vec{r})n(\vec{r}')}{|\vec{r} - \vec{r}'|} + \Delta E_{ee} \quad (34)$$

where the second term obviously comes from the correction in Eq. 33.

The kinetic energy poses an even larger problem. Since the kinetic energy operator contains a derivative term,

$$T = -\frac{1}{2} \int d\vec{r} \Psi^*(\vec{r}_1, \dots, \vec{r}_N) \nabla^2 \Psi(\vec{r}_1, \dots, \vec{r}_N), \quad (35)$$

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. 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

$$n(\vec{r}) = \sum_n^{N_e} |\phi_n(\vec{r})|^2 \quad (36)$$

These orbitals are called *Kohn-Sham* orbitals and they are initially completely unspecified in much the same way as the orbitals in the Slater determinant in the Hartree-Fock formalism[1]. 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.

Writing the density as we have in Eq. 36 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

$$T = -\frac{1}{2} \sum_n^{N_e} \int d\vec{r} \phi_n^*(\vec{r}) \nabla^2 \phi_n(\vec{r}) d\vec{r} + \Delta T \quad (37)$$

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

$$E = -\frac{1}{2} \sum_n^{N_e} \int d\vec{r} \phi_n^*(\vec{r}) \nabla^2 \phi_n(\vec{r}) d\vec{r} + \int n(\vec{r}) V_{ne}(\vec{r}) d\vec{r} + \frac{1}{2} \int \int d\vec{r} d\vec{r}' \frac{n(\vec{r})n(\vec{r}')}{|\vec{r} - \vec{r}'|} + \Delta T + \Delta E_{ee} \quad (38)$$

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

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, ΔE_{ee} and ΔE_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.

$$E_{xc} = \Delta E_{ee} + \Delta T \quad (39)$$

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

$$E_{xc} = \int d\vec{r} n(\vec{r}) \epsilon_{xc}(n) \quad (40)$$

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

$$E = -\frac{1}{2} \sum_n^{N_e} \int d\vec{r} \phi_n^*(\vec{r}) \nabla^2 \phi_n(\vec{r}) d\vec{r} + \int n(\vec{r}) V_{ne}(\vec{r}) d\vec{r} + \frac{1}{2} \int \int d\vec{r} d\vec{r}' \frac{n(\vec{r}) n(\vec{r}')}{|\vec{r} - \vec{r}'|} + \int d\vec{r} \epsilon_{xc}(n) n(\vec{r}). \quad (41)$$

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.

V. 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

$$E^{(1)} = \langle \Psi^{(1)} | \hat{H}^{(1)} | \Psi^{(1)} \rangle < \langle \Psi^{(2)} | \hat{H}^{(1)} | \Psi^{(2)} \rangle. \quad (42)$$

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. 42 as

$$\langle \Psi^{(2)} | \hat{H}^{(1)} | \Psi^{(2)} \rangle = \langle \Psi^{(2)} | \hat{H}^{(2)} | \Psi^{(2)} \rangle + \int d\vec{r} [V_{ext}^{(1)}(\vec{r}) - V_{ext}^{(2)}(\vec{r})] n_0(\vec{r}) \quad (43)$$

Exchanging labels in Eq. 43, we obtain

$$\langle \Psi^{(1)} | \hat{H}^{(2)} | \Psi^{(1)} \rangle = \langle \Psi^{(1)} | \hat{H}^{(1)} | \Psi^{(1)} \rangle + \int d\vec{r} [V_{ext}^{(2)}(\vec{r}) - V_{ext}^{(1)}(\vec{r})] n_0(\vec{r}) \quad (44)$$

Adding Eq. 43 and Eq. 44, we obtain

$$E^{(1)} + E^{(2)} < E^{(2)} + E^{(1)} \quad (45)$$

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.

$$E[n] = T[n] + E_{int}[n] + \int V_{ext}(\vec{r})n(\vec{r}) + E_{II} \equiv F[n] + \int V_{ext}(\vec{r})n(\vec{r}) + E_{II} \quad (46)$$

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, $n^{(1)}(\vec{r})$,

$$E^{(1)} = E[n^{(1)}] = \langle \Psi^{(1)} | \hat{H}^{(1)} | \Psi^{(1)} \rangle. \quad (47)$$

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

$$E^{(1)} = E[n^{(1)}] = \langle \Psi^{(1)} | \hat{H}^{(1)} | \Psi^{(1)} \rangle < \langle \Psi^{(2)} | \hat{H}^{(1)} | \Psi^{(2)} \rangle = E^{(2)}. \quad (48)$$

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.

A. 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.

1. 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.

2. 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

$$\frac{df(x)}{dx} = \frac{2\pi}{N}n(x) \quad (49)$$

which through integration yields

$$f(x) = \frac{2\pi}{N} \int_{-\infty}^x n(x')dx'. \quad (50)$$

Next define a set of single particle orbitals

$$\phi_k(x) = \sqrt{\frac{n(x)}{N}} e^{i[kf(x)+\phi(x)]} \quad (51)$$

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, Ψ_{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

$$\langle \Psi_{k_1, \dots, k_N} | \hat{n} | \Psi_{k_1, \dots, k_N} \rangle = \sum_{i=1}^N |\phi_k|^2 = n(x) \quad (52)$$

where the last equation has been obtained using Eq. 51.

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

$$\hat{H} = \hat{T} + \hat{V}_{ee} + \sum_i \hat{v}_i \quad (53)$$

$F[n]$ 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

$$Q[n] = \min_{\Psi_n \rightarrow n} \langle \Psi_n | \hat{T} + \hat{V}_{ee} | \Psi_n \rangle \quad (54)$$

where a *constrained minimization* is conducted over all the antisymmetric wavefunctions Ψ_n , which yield the trial density n . With the definition in Eq. 54, 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 ,

$$\int d\vec{r} \hat{v}(\vec{r}) n(\vec{r}) + Q[n] \geq E_0 \quad (55)$$

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

$$\hat{V}_{ext} = \sum_i \hat{v}(\vec{r}_i) \quad (56)$$

and E_0 is the ground state energy.

Proof : Following the notation of Levy in his PNAS paper, define Ψ_{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. 54, we have

$$Q[n] = \langle \Psi_{min}^n | \hat{T} + \hat{V}_{ee} | \Psi_{min}^n \rangle \quad (57)$$

and in particular for the ground state density that we seek

$$Q[n_{gs}] = \langle \Psi_{min}^{n_{gs}} | \hat{T} + \hat{V}_{ee} | \Psi_{min}^{n_{gs}} \rangle. \quad (58)$$

The minimum total energy for an arbitrary density, n (the left-hand side of the inequality in Eq. 55) is then given by

$$\int d\vec{r} \hat{v}(\vec{r}) n(\vec{r}) + Q[n] = \int d\vec{r} \hat{v}(\vec{r}) n(\vec{r}) + \langle \Psi_{min}^n | \hat{T} + \hat{V}_{ee} | \Psi_{min}^n \rangle \quad (59)$$

which can alternatively be expressed as

$$\int d\vec{r} \hat{v}(\vec{r}) n(\vec{r}) + Q[n] = \langle \Psi_{min}^n | \hat{T} + \hat{V}_{ee} + \hat{V}_{ext} | \Psi_{min}^n \rangle \quad (60)$$

making use of Eq. 56.

Since, in the general case, the density n in Eq. 60 is not the ground state density we have, by the variational principle,

$$\langle \Psi_{min}^n | \hat{T} + \hat{V}_{ee} + \hat{V}_{ext} | \Psi_{min}^n \rangle \geq E_0. \quad (61)$$

Combining Eq. 60 and Eq. 61, we have

$$\int d\vec{r} \hat{v}(\vec{r}) n(\vec{r}) + Q[n] \geq E_0 \quad (62)$$

which completes the proof of the first theorem.

Theorem II : For the ground state density,

$$\int d\vec{r} \hat{v}(\vec{r}) n_{gs}(\vec{r}) + Q[n_{gs}] = E_0 \quad (63)$$

Proof : Relying on the considerations illustrated so far, the true ground state density of the system Ψ_{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

$$\langle \Psi_{gs} | \hat{T} + \hat{V}_{ee} + \hat{V}_{ext} | \Psi_{gs} \rangle \leq \langle \Psi_{min}^{n_{gs}} | \hat{T} + \hat{V}_{ee} + \hat{V}_{ext} | \Psi_{min}^{n_{gs}} \rangle. \quad (64)$$

Eq. 64 can be rewritten by separating the external potential energy term

$$\int d\vec{r} |\Psi_{gs}|^2 \hat{v}(\vec{r}) + \langle \Psi_{gs} | \hat{T} + \hat{V}_{ee} | \Psi_{gs} \rangle \leq \int d\vec{r} n_{gs}(\vec{r}) \hat{v}(\vec{r}) + \langle \Psi_{min}^{n_{gs}} | \hat{T} + \hat{V}_{ee} | \Psi_{min}^{n_{gs}} \rangle \quad (65)$$

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

$$\langle \Psi_{gs} | \hat{T} + \hat{V}_{ee} | \Psi_{gs} \rangle \leq \langle \Psi_{min}^{n_{gs}} | \hat{T} + \hat{V}_{ee} | \Psi_{min}^{n_{gs}} \rangle. \quad (66)$$

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

$$\langle \Psi_{gs} | \hat{T} + \hat{V}_{ee} | \Psi_{gs} \rangle \geq \langle \Psi_{min}^{n_{gs}} | \hat{T} + \hat{V}_{ee} | \Psi_{min}^{n_{gs}} \rangle. \quad (67)$$

The only way Eq. 66 and Eq. 67 can both be satisfied is if equality holds. In other words,

$$\langle \Psi_{gs} | \hat{T} + \hat{V}_{ee} | \Psi_{gs} \rangle = Q[n_{gs}]. \quad (68)$$

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

$$E_0 = \int d\vec{r} \hat{v}(\vec{r}) n_{gs}(\vec{r}) + \langle \Psi_{gs} | \hat{T} + \hat{V}_{ee} | \Psi_{gs} \rangle \quad (69)$$

Eq. 68 yields

$$\int d\vec{r} \hat{v}(\vec{r}) n_{gs}(\vec{r}) + Q[n_{gs}] = E_0 \quad (70)$$

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. 53 is equal to Eq. 54 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.

3. Fractional Total Particle Number, Discontinuous Derivatives

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

$$\int d\vec{r}n(\vec{r}) = N. \quad (71)$$

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

$$\frac{\delta [E_v[n] - \mu \int d\vec{r}n(\vec{r})]}{\delta n} = 0 \quad (72)$$

where the subscript v denotes an energy that is related to the external potential constructed through \hat{v} and μ 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. 71 can be rewritten as (with the notation taken from Dreizler and Gross, *Density Functional Theory : An Approach to the Quantum Many-Body Problem*)

$$\int d\vec{r}n(\vec{r}) = M + \omega \quad (73)$$

where M is an integer and ω 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

$$F[n] = \min_{n_{\text{mix}}} \langle \Psi_{\text{mix}} | \hat{T} + \hat{V}_{ee} | \Psi_{\text{mix}} \rangle \quad (74)$$

where the minimization goes over all possible mixtures

$$n_{\text{mix}} = \alpha_M \langle \Psi_M | \hat{n}(\vec{r}) | \Psi_M \rangle + \alpha_{M+1} \langle \Psi_{M+1} | \hat{n}(\vec{r}) | \Psi_{M+1} \rangle. \quad (75)$$

Integrating n_{mix} over space and comparing to Eq. 73, we arrive at the rather obvious result

$$\alpha_{M+1} = \omega \quad \text{and} \quad \alpha_M = (1 - \omega). \quad (76)$$

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

$$\frac{\partial E_v[n]}{\partial n} = \mu. \quad (77)$$

As is mostly the case when using Lagrange multipliers, we now try to see if μ 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.

$$\frac{\partial E_N}{\partial N} = \lim_{\epsilon \rightarrow 0} \frac{E_{N+\epsilon} - E_N}{\epsilon} = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \int d\vec{r} \frac{\delta E_v[n]}{\delta n(\vec{r})} \Big|_{n_N} (n_{N+\epsilon} - n_N) = \frac{1}{\epsilon} \mu \int (n_{N+\epsilon} - n_N) d\vec{r} = \frac{1}{\epsilon} \mu \epsilon = \mu \quad (78)$$

where the last equations have been obtained by substituting Eq. 77 into Eq. 78. 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

$$E_{\text{tot}} = E(X) + E(Y). \quad (79)$$

Now assume that $\mu(X) < \mu(Y)$. If we now transfer an infinitesimal amount of charge, δN_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. 74, the minimum energy for a fractional number of particles is (in the Lieb-Levy formalism)

$$E_{M+\omega} = \min_{n(\vec{r}) \rightarrow N+\omega} \min_{\Psi_M, \Psi_{M+1} \rightarrow n(\vec{r})} \left[(1 - \omega) \langle \Psi_M | \hat{H} | \Psi_M \rangle + \omega \langle \Psi_{M+1} | \hat{H} | \Psi_{M+1} \rangle \right] \quad (80)$$

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

$$E_{M+\omega} = (1 - \omega)E_M + \omega E_{M+1}. \quad (81)$$

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. 77. 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

$$I(Z) = E_{Z-1}(Z) - E_Z(Z) \quad (82)$$

$$A(Z) = E_Z(Z) - E_Z(Z + 1) \quad (83)$$

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

$$\begin{aligned} \Delta E_{Y \rightarrow X} &= \underbrace{E_{X+1}(X) + E_Y(Y - 1)}_{\text{final state}} - \underbrace{E_X(X) + E_Y(Y)}_{\text{initial state}} \\ &= E_{X+1}(X) - E_X(X) + E_{Y-1}(Y) - E_Y(Y) = -A(X) + I(Y). \end{aligned} \quad (84)$$

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

$$\Delta E_{X \rightarrow Y} = -A(Y) + I(X). \quad (85)$$

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.

VI. 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

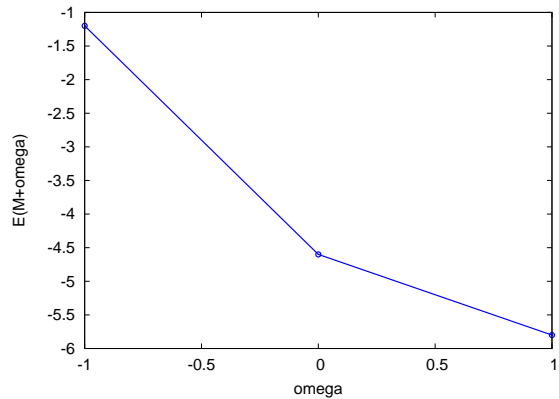
$$\frac{\delta E_e}{\delta \phi_i^*(\vec{r})} = \frac{\delta T_s}{\delta \phi_i^*(\vec{r})} + \left[\frac{\delta E_{ext}}{\delta n(\vec{r})} + \frac{\delta E_{Hartree}}{\delta n(\vec{r})} + \frac{\delta E_{xc}}{\delta n(\vec{r})} \right] \frac{\delta n(\vec{r})}{\delta \phi_i^*(\vec{r})} = \varepsilon_i \phi_i(\vec{r}) \quad (86)$$

$$-\frac{1}{2} \nabla^2 \phi_i(\vec{r}) + \left[V_{ext}(\vec{r}) + \underbrace{\int d\vec{r}' \frac{n(\vec{r}')}{|\vec{r}' - \vec{r}'|}}_{V_H} + \underbrace{\epsilon_{xc}[n] + n(\vec{r}) \frac{\delta \epsilon_{xc}[n]}{\delta n(\vec{r})}}_{V_{xc}} \right] \phi_i(\vec{r}) = \varepsilon_i \phi_i(\vec{r}) \quad (87)$$

Eq. 87 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

$$[\hat{T} + V_{eff}] \phi_i(\vec{r}) = \varepsilon_i \phi_i(\vec{r}), \quad (88)$$

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.



A. Issues regarding the Kohn-Sham methods

1. 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. 87 from the left by $\phi^*(\vec{r})$, integrating over space and summing over i . This sequence of operations yield

$$\sum \epsilon_i = T_s + \int d\vec{r} n(\vec{r}) V_{ext} + \int \int d\vec{r} d\vec{r}' \frac{n(\vec{r})n(\vec{r}')}{|\vec{r} - \vec{r}'|} + \int d\vec{r} n(\vec{r}) \epsilon_{xc}(n(\vec{r})) + \int d\vec{r} n^2(\vec{r}) \frac{\partial \epsilon_{xc}}{\partial n} \quad (89)$$

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

$$\Delta E = \sum_i \epsilon_i - E_{tot}[n] = \frac{1}{2} E_H + \int d\vec{r} n^2(\vec{r}) \frac{\partial \epsilon_{xc}}{\partial n}. \quad (90)$$

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).

2. 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, ϵ_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

$$n(\vec{r}) = \sum_{i=1}^N |\psi_i|^2 = \sum_{i=1}^{\infty} n_i |\psi_i|^2 \quad (91)$$

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

$$\sum_{i=1}^{\infty} n_i = N \quad \text{with} \quad 0 \leq n_i \leq 1. \quad (92)$$

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

$$n_i = \cos^2 \theta_i. \quad (93)$$

The generalized energy is now

$$\tilde{E}[\psi_1, \dots; \gamma_i, \dots] = \sum_{i=1}^{\infty} n_i t_i + \int n(\vec{r}) V_{eff}(\vec{r}) d\vec{r} \quad (94)$$

with the above definition of $n(\vec{r})$ in Eq. 91. In Eq. 94, we have employed Janak's notation (modified slightly)

$$t_i = -\frac{1}{2} \int \psi_i^*(\vec{r}) \nabla^2 \psi_i(\vec{r}) d\vec{r} = \epsilon_i - \int \psi_i^*(\vec{r}) (V_H + V_{xc} + v_{ext}(\vec{r})) \psi_i(\vec{r}) d\vec{r} \quad (95)$$

where the latter equality follows straightforwardly from Eq. 87.

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)

$$Q[\psi_1, \dots; \theta_1, \dots] = \tilde{E}[\psi_1, \dots; \theta_1, \dots] - \sum_i \lambda_i \left[\int |\psi_i|^2 d\vec{r} - 1 \right] - \mu \left[\sum_{i=1}^{\infty} n_i - N \right] \quad (96)$$

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 θ_i gives us

$$(\sin 2\theta_i) [\langle \psi_i | \hat{h} | \psi_i \rangle - \mu] = (\sin 2\theta_i) [\epsilon_i - \mu] = 0 \quad (97)$$

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

$$\begin{aligned} \epsilon_i = \mu \quad \text{and} \quad \theta_i \text{ is arbitrary} &\Rightarrow 0 \leq n_i \leq 1 \\ \epsilon_i \neq \mu \quad \text{and} \quad \theta_i \text{ is arbitrary} &\Rightarrow n_i = 1 \\ \epsilon_i \neq \mu \quad \text{and} \quad \theta_i \text{ is arbitrary} &\Rightarrow n_i = 0 \end{aligned} \quad (98)$$

From Eq. 98, 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. 94 gives

$$\frac{\partial \tilde{E}}{\partial n_i} = t_i + \int d\vec{r} V_{eff} |\psi_i|^2 = \epsilon_i, \quad (99)$$

which is the common statement of Janak's theorem. Note that in Janak's original work, the orbitals ψ_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

$$E_{N+1} - E_N = \int_N^{N+1} dn_M \frac{\partial E_M}{\partial n_M} = \int_N^{N+1} dn_M \epsilon_M \quad (100)$$

where ϵ_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.

Hartree (1 H=27.211 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

$$\hat{H} = -\frac{1}{2} \sum_i^{N_e} \nabla_i^2 - \sum_i^{N_e} \sum_I^{N_n} \frac{Z_I}{|\vec{r}_i - \vec{R}_I|} + \frac{1}{2} \sum_i^{N_e} \sum_{j \neq i}^{N_e} \frac{1}{|\vec{r}_i - \vec{r}_j|} + \frac{1}{2} \sum_I^{N_n} \sum_{J \neq I}^{N_n} \frac{Z_I Z_J}{|\vec{R}_I - \vec{R}_J|}. \quad (101)$$

[1] One can prove (by example) that it is always possible to construct such orbitals and that they are not unique. In fact, one can construct infinitely many such orbitals