

# 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 :

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

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 :

$$\begin{aligned} f(x) : x &\rightarrow y \quad \text{function} \\ F[f(x)] : f(x) &\rightarrow y \quad \text{functional} \end{aligned} \quad (12)$$

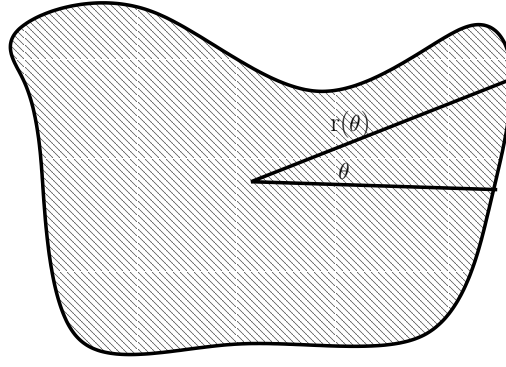
As seen in Eq. 12, arguments of functionals are written in square brackets to distinguish them from regular functions. The perimeter and area of a closed curve parametrized by an angle are familiar examples of functionals.

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

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

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