Notes on the *ab initio* theory of molecules and solids: Density functional theory (DFT)

Tomás A. Arias

February 2, 2000

Cornell University

Department of Physics

Contents

1 Introduction **2**

2 The importance of the total energy **2**

3 The total energy within density functional theory **2**
3.1 Electrostatics . . . 3
3.2 Quantum mechanics . . . 4
3.3 Variational principle . . . 5

4 Kohn-Sham equations **6**
4.1 Basics of the calculus of variations . . . 6
4.2 Derivative of a real function of a complex variable and its conjugate . . . 7
4.3 Kohn-Sham Equations . . . 7
4.4 Solution of the equations . . . 9

5 Atomic units **10**

6 References **11**

1 Introduction

Dating back to at least the time of the ancient philosopher Empedocles (born ~450 BCE) and his theory of the elements earth, water, air and fire, and their basic interactions of love and strife, humanity has striven to understand the behavior of the material world "from the beginning," *ab initio*.

It has taken twenty-three centuries to bring this dream to fruition. C. Coulomb gave us the modern understanding of what what would prove the basic interaction, electrostatics, in the late 1780's. It would take another hundred years to identify the basic constituents of matter as electrons (J.J. Thompson in 1897 and R.A. Millikan in 1909) and nuclei (E. Rutherford in 1911), and yet another two decades to formulate the final ingredient needed for a predictive theory of these tiny objects, quantum mechanics. The rapid, heady developments of early twentieth century prompted P.A.M. Dirac in 1929 to make the following statement of optimism tempered with disappointment:

> The general theory of quantum mechanics is now almost complete. The underlying physical laws ... for ... a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that ... these laws lead to equations much to difficult to be solvable.

Despite the amazingly rapid progress of the early twentieth century, it would take nearly another fifty years to surmount the difficulties which Dirac foresaw. However, it is now possible for us to write software running on a personal computer to solve these equations. A combinations of three developments makes this possible: (1) the development of density functional theory (DFT), for which Walter Kohn shared the 1998 Nobel prize in Chemistry and which is the subject of these notes; (2) the development of powerful new numerical methods, which are the main subject of this course; and (3) the exponential progress in computer power. In this course you will exploit these three developments to fulfill Empedocles' dream for yourself.

2 The importance of the total energy

Figure 1 illustrates the modern scientific view of a typical material system such as a molecule or solid. The nuclei are small, point-like objects, and we denote their locations in $d = 3$ dimensional space as $\vec{X}_1$, $\vec{X}_2$, ... The electrons, which have much smaller mass, move much more quickly and spread out into "electron clouds" which we learn below to describe in terms of orbitals.

The large difference in the mass of the nuclei and electrons allows us to consider the nuclei as nearly stationary compared to the electrons so that we may consider total energy of the system $E(\vec{X}_1, \vec{X}_2, \ldots)$ for each particular arrangement of the nuclei, a quantity which gives us a tremendous knowledge of the system. For instance, the minimum value of E over all possible nuclear positions $\vec{X}_1, \vec{X}_2, \ldots$ gives the binding energy of the system. The arrangement of the nuclei for which the energy is minimized gives the equilibrium structure, including all of the bond lengths, bond angles and, for solids, the lattice parameters. The derivatives of E with respect to the coordinates of the nuclei give the forces $\vec{F}_i = -\nabla_{\vec{X}_i} E$ on the nuclei, which determine how the system evolves in time. From this, we can compute the vibrational frequencies and even the melting point of the system. These are just a few of the things which we can compute *ab initio*, using only a few fundamental constants of nature as experimental input!

3 The total energy within density functional theory

Density functional theory (developed by Hohenberg, Kohn, and Sham [1,2]) gives a relatively simple prescription for the total energy of a system of electrons and nuclei. For the purpose of this course, you may take this discussion as giving our basic postulates of quantum mechanics. (For a more fundamental discussion, you may wish to consider Phys 683.)

There are two basic types of energy which we must consider, potential energy and the kinetic of the electrons. (Remember that we regard the nuclei as stationary.) We begin below in Section 3.1 with the potential energy, which classical electrostatic theory describes quite well. Then, to deal with the kinetic energy, we will have to introduce quantum mechanics in Section 3.2.

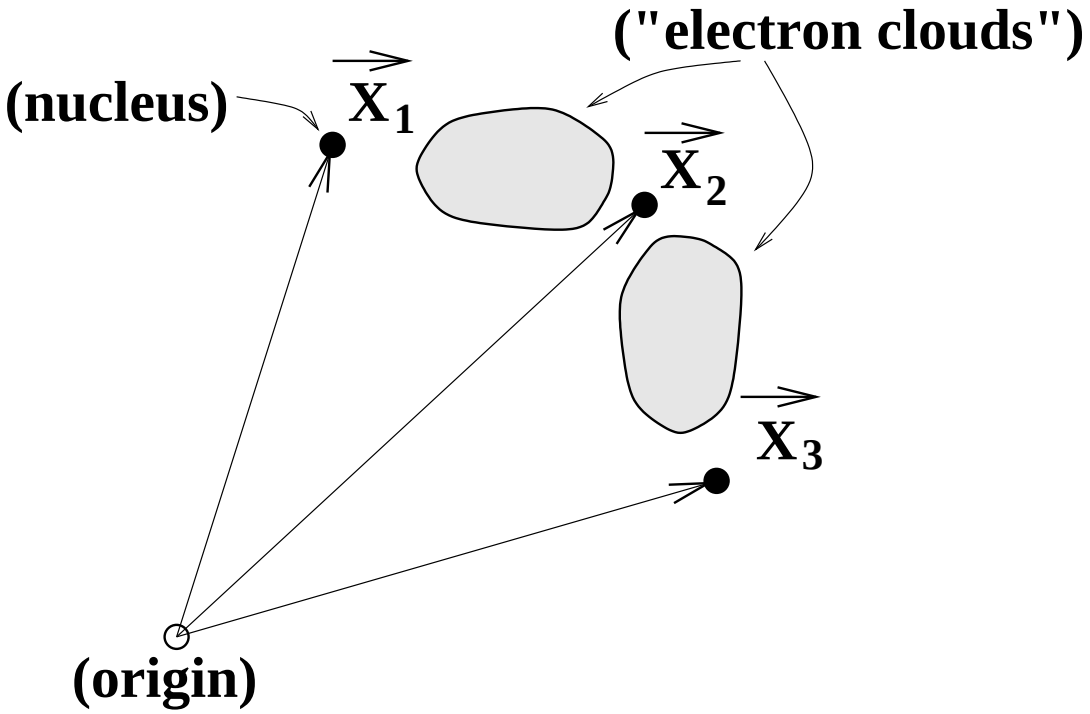


Figure 1: Modern view of a molecular or solid system: nuclei (large dots), electron clouds (grey shaded regions).

3.1 Electrostatics

There are three groups of electrostatic interactions which we must consider: interactions of nuclei with nuclei, of electrons with nuclei, and of electrons with electrons. Coulomb's law states that the potential energy between two charges q_1 and q_2 at separation r_{12} is

$$U = [k_c]\frac{q_1 q_2}{r_{12}},$$

where $[k_c]$ is Coulomb's constant[1].

The total potential energy arising from the interactions among all the nuclei is just the sum of all pair-wise interactions,

$$U_{nuc-nuc} = \frac{1}{2}[k_c]e^2 \sum_{I \neq J} \frac{Z_I Z_J}{R_{IJ}}, \tag{1}$$

where e is the charge of the electron, I and J index different nuclei, R_{IJ} is the separation between nuclei I and J, Z_I is the atomic number of nucleus I, and the factor of $1/2$ is the famous double-counting correction to ensure that we count each pair-wise interaction only once.

Similarly, the potential energy of a *single* electron at position $\vec{x}$ due to the nuclei is

$$V_{nuc}(\vec{x}) = -[k_c]e^2 \sum_i \frac{Z_I}{R_I}, \tag{2}$$

where R_I is the distance from point $\vec{x}$ to nucleus I. If the volume density (number per unit volume) of electrons is $n(\vec{x})$, then the number of electrons in the volume element dV near point $\vec{x}$ is $n(\vec{x})\,dV$ and the

[1] Note that in the cgs systems of units $k_c \equiv 1$. If you are more comfortable working in such units, then simply ignore any factors which appear below in square brackets.

total potential energy of the electrons interacting with the nuclei is

$$U_{el-nuc} = \int V_{nuc}(\vec{x}) n(\vec{x})\, dV, \tag{3}$$

where the integral is over all of space. (There is no 1/2 double-counting correction here because the interaction between electron #1 and nucleus #2 is not counted again when we do the interaction between electron #2 and nucleus #1!)

Finally, the electrons interact not only with the nuclei, but also with themselves. From Coulomb's law, the potential energy for a single electron at point $\vec{x}$ coming from the electrons at point $\vec{x}'$ is $[k_c]e^2\, n(\vec{x}')\, dV'/|\vec{x}-\vec{x}'|$ where $|\vec{x}-\vec{x}'|$ is the distance between points $\vec{x}$ and $\vec{x}'$. The total potential for a single electron at point $\vec{x}$ is then

$$\phi(\vec{x}) = [k_c]e^2 \int \frac{n(\vec{x}')\, dV'}{|\vec{x}-\vec{x}'|}.$$

Standard electrostatics tells us that doing this integral is equivalent to solving Poisson's equation,

$$\nabla^2 \phi(\vec{x}) = -4\pi [k_c] e^2 n(\vec{x}). \tag{4}$$

Finally, once we have $\phi(\vec{x})$, the potential energy for the electrons interacting with themselves follows the same logic as (3) but with the double-counting correction of (1) because we are dealing with the total interaction of a group of particles with itself,

$$U_{el-el} = \frac{1}{2} \int \phi(\vec{x}) n(\vec{x})\, dV. \tag{5}$$

3.2 Quantum mechanics

Eqs. (1,3,5) describe the potential energy of the system, but we yet have to determine the electron density $n(\vec{x})$ and have yet to consider the kinetic energy of the electrons. Density functional theory determines both of these quantities.

Within density functional theory a set of quantum mechanical *Kohn-Sham orbitals* $\psi_i(\vec{x})$ describes the electrons. These are the electronic orbitals that you learn about in introductory chemistry class, each of which usually contains two electrons (one spin-up and one spin-down). In general, these orbitals may be complex, so that we must also consider the complex conjugates of the orbitals, $\psi_i^*(\vec{x})$. For the problems of interest in their course, the orbitals always turn out to be real so that $\psi_i^*(\vec{x}) = \psi_i^*(\vec{x})$. Thus, if you are unfamiliar or rusty with complex numbers you can simply ignore the *'s. We include them for those in the course who are familiar with quantum mechanics and who may have interest in problems where the orbitals can be complex.

Within quantum mechanics, the square magnitude of each orbital gives the probability of finding an electron, when in that orbital, at any point in space, $\mathcal{P}(\vec{x}) = \psi_i^*(\vec{x})\psi_i(\vec{x}) \equiv |\psi_i(\vec{x})|^2$. The orbitals are not free to be any functions whatsoever, but must obey certain constraints. First, because the electron must be somewhere in space, the probability must add up to unity,

$$1 = \int \mathcal{P}(\vec{x})\, dV = \int \psi_i^*(\vec{x})\psi_i(\vec{x})\, dV. \tag{6}$$

In addition to this *normality* constraint for each orbital i, the orbitals must be *orthogonal* to each other,

$$0 = \int \psi_i^*(\vec{x})\psi_j(\vec{x})\, dV \quad \text{for } i \neq j, \tag{7}$$

the condition by which density functional theory encodes the *Pauli exclusion principle* from elementary chemistry courses. Apart from these constraints, the orbitals are completely free. Thus, we may combine all relevant constraints into the *orthonormality* constraint,

$$\int \psi_i^*(\vec{x})\psi_j(\vec{x})\, dV = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}. \tag{8}$$

The electron density and total kinetic energy come directly from the orbitals. Because the square of each orbital gives the distribution of the electrons in that orbital, the total electron density will be the sum of squares of the orbitals time the number of electrons f_i in or "filling" each orbital,

$$n(\vec{x}) = \sum_i f_i \, |\psi_i(\vec{x})|^2. \tag{9}$$

(As mentioned above, usually there are $f_i = 2$ electrons in each orbital.) The total kinetic energy of the electrons T_{el} is similarly just the sum over orbitals of the number of electrons in each orbital times the elementary quantum mechanical expression for the kinetic energy of each orbital,

$$T_{el} = \sum_i f_i \int \psi_i^*(\vec{x}) \left(-\frac{\hbar^2}{2m} \nabla^2 \psi_i(\vec{x}) \right) \, dV. \tag{10}$$

There arises from advanced quantum mechanics one final subtle point. The electron density defined in (9) is only an average. The actual density fluctuates, resulting in relatively small but important errors in Eqs. (5,10) due to correlations in these fluctuations. In theory, we may correct for these errors exactly, but this turns out to be quite difficult in practice. A very good approximation to this *exchange-correlation* correction, sufficient in practice to compute most properties to within a few percent, is the *local density approximation*

$$E_{xc} = \int f_{xc}(n(\vec{x})) \, dV, \tag{11}$$

where $f_{xc}(\ldots)$ is a relatively simply function which we will provide later in the course[2].

That's it – this is all the quantum mechanics we need to predict accurately the behavior of matter!

3.3 Variational principle

Putting everything together, we now have our expression for the total energy,

$$\begin{aligned} E[\{\psi(\vec{x})\}] \;\; = \;\; & \sum_i f_i \int \psi_i^*(\vec{x}) \left(-\frac{\hbar^2}{2m} \nabla^2 \psi_i(\vec{x}) \right) \, dV + \int V_{nuc}(\vec{x}) n(\vec{x}) \, dV \\ & + \frac{1}{2} \int \phi(\vec{x}) n(\vec{x}) \, dV + \int f_{xc}(n(\vec{x})) \, dV + U_{nuc-nuc}, \end{aligned} \tag{12}$$

where $f_{xc}(\ldots)$ is some known function, $V_{nuc}(\vec{x})$ is the potential energy field created by the nuclei, $U_{nuc-nuc}$ is the simple electron static interaction among the nuclei,

$$n(\vec{x}) = \sum_i f_i \, |\psi_i(\vec{x})|^2,$$

and f_i (usually equal to two) is the number of electrons in orbital i. Note that the expression (12) maps each possible choice of the set of electronic orbitals $\{\psi_i(\vec{x})\}$ to a unique value for the energy of the system and thereby gives the total energy E as a function of the orbital functions $\phi_i(\vec{x})$. Such an expression which returns a number as a function of other functions is called a function*al* and denoted with square brackets as we do in Eq. (12).

We now have a functional for the energy in terms of the orbitals, but which orbitals are the right ones to use? The answer is quite sensible: the correct orbitals are those which minimize the total energy E in (12) while obeying the orthonormality constraints (6). Combined with this *variational principle*, Eq. (12) now gives a complete prescription for computing total energies, and thereby all of the properties mentioned in Sec. 2.

[2] Improving the approximations for E_{xc} is one of the "holly grails" of electronic structure. If this interests you, please let me know ...

4 Kohn-Sham equations

There are two schools of thought on how to achieve the minimization of the total energy. The more prevalent approach in the physics community is to view the calculation directly as a problem in numerical minimization and to apply modern techniques for constrained numerical minimization. We shall return to this approach in the second half of the this course. The second school of thought, more prevalent in the chemistry community, is to derive the Lagrange-multiplier equations for constrained minimization and to then use numerical methods to solve the resulting equations. As we shall see, each approach has its advantages and disadvantages. In the end, though, both must lead to the same result.

We now derive Lagrange-multiplier equations for density functional theory, known as the Kohn-Sham equations.

4.1 Basics of the calculus of variations

To derive the Kohn-Sham equations, we must first take the derivative of a functional, which is the basic subject of the *calculus of variations*. Despite the mystique associated with the calculus of variations, it is really no more complicated than taking derivatives of multi-variable functions. This is because any functional, say $F[g(x)]$, may be viewed as just a function of a large collection of variables, namely the values of its argument function $g(x)$ at each point in space x. One can think of a function $g(x)$ as a (very long) vector of values, one for each value of x, just as we think of a vector $\vec{q}$ as a set of values q_i, one for each value of the index i. With this perspective, the points in space x are the analogue of the index i, so that we can think of the function $g(x)$ also as the indexed set of values g_x.

With this perspective, we see that, just as the first-order variation df of a multi-variable function $f(\vec{q})$ with changes in its argument $\vec{q}$ is given by a sum over the index i

$$\begin{aligned} \delta f &\equiv f(\vec{q}+\delta\vec{q}) - f(\vec{q}) \\ &= (\nabla f(\vec{q}))\cdot\delta q \quad \text{(to first order)} \\ &= \sum_i \left(\frac{\partial f}{\partial q_i}\right)\delta q_i, \end{aligned}$$

the variation of the functional $F[g(x)]$ is given by a "sum" over the index x,

$$\begin{aligned} \delta F &\equiv F[g(x)+\delta g(x)] - F[g(x)] \\ &= \int \left(\frac{\delta F}{\delta g(x)}\right)\delta g(x)\; dx. \end{aligned} \tag{13}$$

Note that, because x is now a continuous variable, the "sum" becomes an integral. also, $(\delta F/\delta g(x))$ is the standard notation for the *functional derivative*, which we see amounts to taking the partial derivative of F with respect to the value $g(x)$. With this understood, we can take functional derivatives as easily as differentiating a multi-variable function. All of the usual rules still apply, such as the product and chain rules!

As an example, let us consider the functional derivative of $E_{xc}[n(\vec{x})]$ with respect to $n(\vec{x})$. First, we shall carry out the variation formally, and then we shall show how quickly we arrive at the same result by analogy with multi-variable calculus. Applying the formal definition (13) of the functional derivative to E_{xc} in (11), we find

$$\begin{aligned} \delta E_{xc} &\equiv E_{xc}[n(\vec{x})+\delta n(\vec{x})] - E_{xc}[n(\vec{x})] \\ &= \int f_{xc}\left(n(\vec{x})+\delta n(\vec{x})\right)\, dV - \int f_{xc}\left(n(\vec{x})\right)\, dV \\ &= \int \left(f_{xc}\left(n(\vec{x})+\delta n(\vec{x})\right) - f_{xc}\left(n(\vec{x})\right)\right)\, dV \\ &= \int f'_{xc}\left(n(\vec{x})\right)\,\delta n(\vec{x})\, dV \quad \text{(to first order)}, \end{aligned}$$

from which we may read off the result

$$\frac{\delta E_{xc}}{\delta n(\vec{x})} = f'_{xc}(n(\vec{x})). \tag{14}$$

Alternatively, we note that Eq. (11) is just the integral of the result of applying the function $f_{xc}(\ldots)$ to each component of $n(\vec{x})$ separately. If this were a multi-variable problem, the analogous function would be a sum over the values of a function evaluated separately on each component,

$$e_{xc}(\vec{q}) = \sum_i f_{xc}(q_i),$$

and the derivatives are then

$$\frac{\partial e_{xc}}{\partial q_i} = f'_{xc}(q_i)$$

because only the "i" term in the sum depends on q_i. Changing the ∂ symbols to δ, replacing q with n and the index i with $\vec{x}$, we arrive immediately at precisely (14)!

4.2 Derivative of a real function of a complex variable and its conjugate

Despite the fact that the orbitals $\psi_i(\vec{x})$ may be complex, the energy function (12) always turns out to be real. We can use this to take a very powerful short cut which is often used but infrequently explained.

For simplicity of notation, let us consider minimizing a real function $f(\ldots)$ of a single complex variable z. One way to think of this problem as minimizing a real function of two independent real variables, namely the real and imaginary parts of $z \equiv z_r + z_i\, i$. To minimize, we then need to compute the two partial derivatives $\partial f(z_r, z_i)/\partial z_r|_{z_i}$ and $\partial f(z_r, z_i)/\partial z_r|_{z_r}$, which will be real because f is always real.

Frequently, however, we are given the function f not in terms of z_r and z_i, but in terms of z and z^*. Noting that $z_r = (z + z^*)/2$ and $z_i = (z - z^*)/2i$, we may use the chain rule to take the derivative of $f(z, z^*)$ with respect to z^* while treating z as a constant,

$$\begin{aligned} \left.\frac{\partial f}{\partial z^*}\right|_z &= \left.\frac{\partial z_r}{\partial z^*}\right|_z \left.\frac{\partial f}{\partial z_r}\right|_{z_i} + \left.\frac{\partial z_i}{\partial z^*}\right|_z \left.\frac{\partial f}{\partial z_i}\right|_{z_r} \\ &= \frac{1}{2}\left(\left.\frac{\partial f}{\partial z_r}\right|_{z_i} + i\left.\frac{\partial f}{\partial z_i}\right|_{z_r}\right). \end{aligned} \tag{15}$$

Thus, the real and imaginary components of $\partial f(z, z*)/\partial z^*|_z$ give us both derivatives $\partial f(z_r, z_i)/\partial z_r$ and $\partial f(z_r, z_i)/\partial z_r$ *simultaneously*. In particular, to minimize over all possible values of $z = z_r + iz_i$, we need just one equation!

$$0 = \left.\frac{\partial f}{\partial z^*}\right|_z.$$

4.3 Kohn-Sham Equations

Using what we have just learned about taking derivatives of real functions of complex variables and including the normality constraint of each orbital $\psi_i(\vec{x})$ with a separate Lagrange multiplier λ_i, the condition for the constrained minimization of (12) is

$$0 = \frac{\delta}{\delta\psi_i^*(\vec{x})}\left(T_{el} + U_{el-nuc} + U_{el-el} + E_{xc} + U_{nuc-nuc} - \sum_i \lambda_i \int \psi_i^*(\vec{x})\psi_i(\vec{x})\, dV\right).$$

We now take this derivative term by term.

The kinetic energy (10) has only one $\psi_i^*(\vec{x})$ term in it, so the derivative is just what this term multiplies,

$$\frac{\delta}{\delta\psi_i^*(\vec{x})}(T_{el}) = -f_i\,\frac{\hbar^2}{2m}\psi_i(\vec{x}).$$

For the electron-nuclear potential energy (3), the only term which depends on $\psi_i^*(\vec{x})$ is the charge density, where $\psi_i^*(\vec{x})$ multiplies $f_i\,\psi_i(\vec{x})$. The nuclear potential $V_{nuc}(\vec{x})$ is unchanged as $\psi_i^*(\vec{x})$ varies, so the final term is just

$$\frac{\delta}{\delta\psi_i^*(\vec{x})}\,(U_{el-nuc}) = f_i\,V_{nuc}(\vec{x})\psi_i(\vec{x}).$$

The electron-electron energy has a very similar structure. The only difference is that, here, when we change $\psi_i^*(\vec{x})$, the potential function $\phi(\vec{x})$ also changes because it depends on $n(\vec{x})$. The net effect of the change in $\phi(\vec{x})$ is the same as that of the direct change in $n(\vec{x})$. To see this we note that Poisson's equation $\nabla^2\phi = -4\pi[k_c]e^2 n$ implies also that $\nabla^2(\delta\phi) = -4\pi[k_c]e^2(\delta n)$. Thus,

$$\int(\delta\phi)n\,dV = \int(\delta\phi)\frac{\nabla^2\phi}{-4\pi[k_c]e^2}\,dV = \int\nabla^2\left(\frac{\delta\phi}{-4\pi[k_c]e^2}\right)\phi\,dV = \int(\delta n)\phi\,dV,$$

where we have moved the ∇^2 from acting on ϕ to acting on $\delta\phi$ by integrating by parts twice. Since both terms are equal, we can take just twice the $(1/2)\int\phi\,\delta n\,dV$ term,

$$\frac{\delta}{\delta\psi_i^*(\vec{x})}\,(U_{el-el}) = f_i\,\phi(\vec{x})\psi_i(\vec{x}).$$

For the exchange-correlation term, we have already derived in (14) that $\delta E_{xc}/\delta n(\vec{x}) = f'_{xc}(n(\vec{x}))$. By the chain rule we just need to multiply this by $\delta n(\vec{x})/\delta\psi_i^*(\vec{x}) = f_i\,\psi_i(\vec{x})$ for the result

$$\frac{\delta}{\delta\psi_i^*(\vec{x})}\,(E_{xc}) = f_i\,f'_{xc}(n(\vec{x}))\psi_i(\vec{x}).$$

Fortunately, $U_{nuc-nuc}$ depends only on the nuclear positions and does not change with $\psi_i^*(\vec{x})$, so

$$\frac{\delta}{\delta\psi_i^*(\vec{x})}\,(U_{nuc-nuc}) = 0.$$

And, finally, $\psi_i^*(\vec{x})$ only appears once in the constraint term, making the derivative,

$$\frac{\delta}{\delta\psi_i^*(\vec{x})}\left(-\sum_i\lambda_i\int\psi_i^*(\vec{x})\psi_i(\vec{x})\,dV\right) = -\lambda_i\,\psi_i(\vec{x}).$$

Summing all of these contributions, setting the resulting equation to zero, moving the "$\lambda_i\psi_i(\vec{x})$" term to the right-hand side, and dividing through by f_i, we get the final result,

$$-\frac{\hbar^2}{2m}\nabla^2\psi_i(\vec{x}) + [V_{nuc}(\vec{x}) + \phi(\vec{x}) + f'_{xc}(n(\vec{x}))]\,\psi_i(\vec{x}) = \frac{\lambda_i}{f_i}\psi_i(\vec{x}).$$

Fortunately for us, this is in the form of a very well-known equation for which there are standard techniques. This is in the form of the standard Schrödinger equation,

$$-\frac{\hbar^2}{2m}\nabla^2\psi_i(\vec{x}) + V(\vec{x})\psi_i(\vec{x}) = \epsilon_i\psi_i(\vec{x}), \tag{16}$$

where we define the potential term as

$$V(\vec{x}) \equiv V_{nuc}(\vec{x}) + \phi(\vec{x}) + f'_{xc}(n(\vec{x})), \tag{17}$$

and we define $\epsilon_i \equiv \lambda_i/f_i$. We interpret the potential $V(\vec{x})$ as just the sum of the nuclear potential, the electrostatic potential $\phi(\vec{x})$ created by the electrons, and an extra, "exchange-correlation" potential correction, $V_{xc}(\vec{x}) \equiv f'_{xc}(n(\vec{x}))$. Since the Lagrange-multipliers are unknown constants at the start, we may as well think in terms of the $\epsilon_i \equiv \lambda_i/f_i$ instead, which have the interpretation of the Schrödinger energies for each orbital.

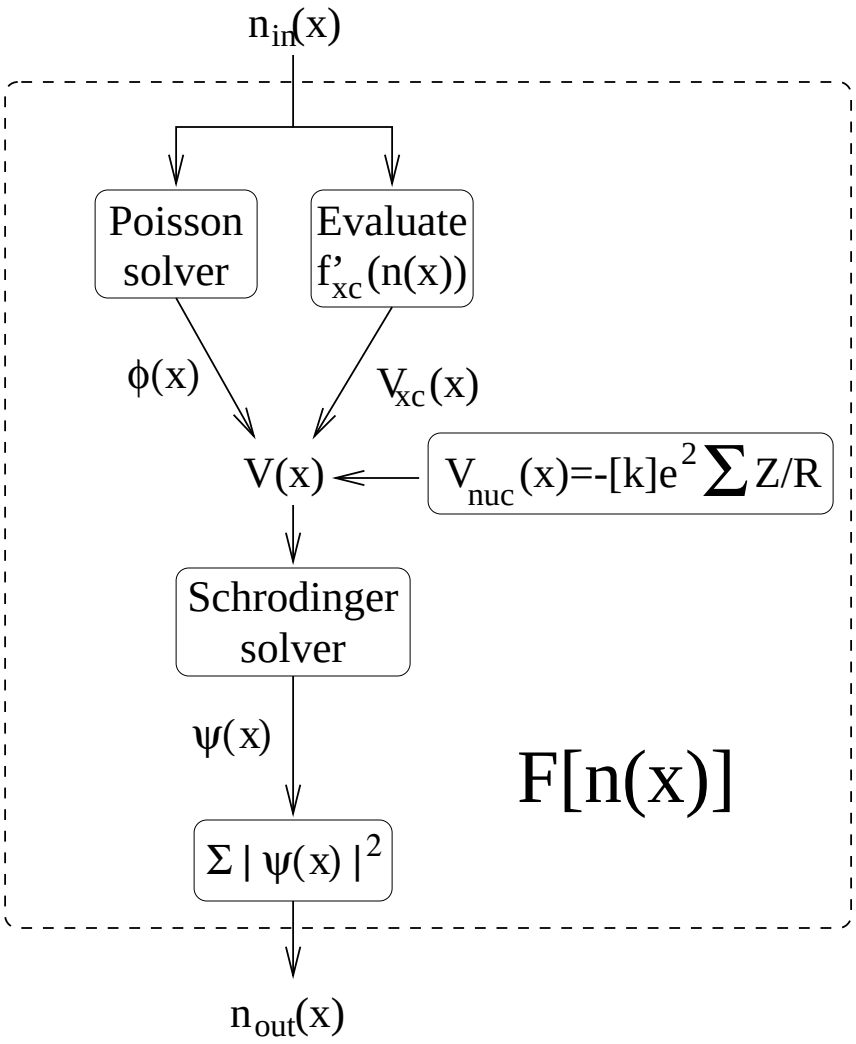


Figure 2: Stages of solving the Kohn-Sham equations

4.4 Solution of the equations

We are now prepared to outline a general strategy for finding the total energy E. To evaluate the various terms in (12), we need the correct orbitals $\psi_i(\vec{x})$, which we can find from $V(\vec{x})$ using a Schrödinger-solver routine to solve (16). To determine $V(\vec{x})$ according to (17), we need the nuclear potential, which we know from (2); the potential from the electrons, which we get from a Poisson-solver routine to solve (4) for a given $n(\vec{x})$; and a simple subroutine to evaluate the function $f'_{xc}(\ldots)$ for the values of $n(\vec{x})$.

The minimum of E obtains only when all of the above equations hold *simultaneously*. In particular, the density must be self-consistent: the density $n_{in}(\vec{x})$ which we input must lead to a potential $V(\vec{x})$ which gives rise to a set of orbitals $\psi(\vec{x})$ that sum to a final density $n_{out}(\vec{x})$ equal to the input density. (See Figure 2.) Viewing the contents of the dashed box in Figure 2 as a function $F[n(\vec{x})]$ which takes the density as input and produces a new density as output, self-consistency is the condition that $F[n(\vec{x})] - n(\vec{x}) = 0$, a set of non-linear equations for the value of the charge density at each point in space. There are quite power numerical methods for solving such equations given the capability of computing $F[n(\vec{x})] - n(\vec{x})$. Eventually, we shall discuss these techniques, but first we shall develop the ability to compute $F[\ldots]$.

5 Atomic units

Because even the simplest mistake can result in hours of debugging, it is critical to do everything possible to make software clean and simple. One thing which we can do in the physical sciences toward this end is to use dimensional analysis.

The simplest form of dimensional analysis is to change to a new system of units tailored specifically to the problem at hand. This is relatively straight-forward because it involves no change in our equations and changes only the numerical values of the physical constants which appear. It requires only that, once the calculations are complete, we convert the results back to standard units using the familiar rules for unit conversion.

A tailored system of units can be very useful when the relevant physical constants have large or small values in term standard units. In a quantum mechanics calculation, for instance, $\hbar \approx 10^{-34}$ kg m^2/s^2 and, as $\hbar^2$ appears in many of our expressions, numerical underflow is a significant risk. On the other hand, if we worked not in meters but Angstroms (1 Å$= 10^{-10}$ m), which are much more relevant for quantum mechanical problems, then we have $\hbar \approx 10^{-14}$ kg Å^2/s^2, a much more manageable number.

Often times, we can do much better and arrange so that all of the relevant physical constants have a numerical value of unity. This has the tremendous advantage that we do not have to type the values of the physical constants into each subroutines or try to set up a repository of global variables, both of which are a frequent source of hard-to-track bugs. Once should definitely seek such an appropriate set of units before beginning a scientific application.

In the case of density functional theory, inspection of the expressions above reveals four physical constants: Planck's constant $\hbar$, the electron mass m, Coulomb's constant $[k_c]$, and the electron charge e. Rather than use the standard units of meter, kilogram and second for the three fundamental dimensions of length, mass and time, we can define three new units, which we shall call L, M and T, respectively. With the ability to choose three unknowns, in general we can hope to reduce only three physical constants to the value unity. We are fortunate, however, because our physical constants always appear in only one of *two* different combinations, $\hbar^2/m$ or $[k_c]e^2$, which we can simultaneously reduce to unity with an appropriate choice of units.

For our two combinations, we have

$$\begin{aligned} \frac{\hbar^2}{m} &= 1.22\,085\,40 \times 10^{-38}\ \mathrm{J\,m^2} \\ [k_c]\,e^2 &= 2.30\,707\,955 \times 10^{-28}\ \mathrm{J\,m}, \end{aligned}$$

where J represents the SI unit of energy, the Joule, which has units 1 J=1 kg m^2/s^2. In our new system of units, we would like these combinations of constants to appear as

$$\begin{aligned} \frac{\hbar^2}{m} &= 1\ \mathrm{E\,L^2} \\ [k_c]\,e^2 &= 1\ \mathrm{E\,L}, \end{aligned} \tag{18}$$

where E is the unit of energy in our units 1 E=1 M L^2/T^2. We may easily solve (18) for L and E, finding the standard atomic units of the Bohr and the Hartree as our units of length and energy, respectively,

$$\mathrm{L} \equiv 1\ \mathrm{bohr} = \frac{\hbar^2/m}{[k^c]e^2} = 0.529\,177\,25 \times 10^{-10}\ \mathrm{m} = 0.529\,177\,25\ \text{Å} \tag{19}$$

$$\mathrm{E} \equiv 1\ \mathrm{hartree} = \frac{([k_c]e^2)^2}{\hbar^2/m} = 4.35\,974\,82 \times 10^{-18}\mathrm{J} = 27.2\,113\,96\mathrm{eV} \tag{20}$$

From now, on so long as we interpret all distances in our calculations as expressed in Bohrs (about 1/2 Angstrom) and all energies in Hartrees (about 27 electron Volts), we can take the factors $\hbar^2/m$ and $[k_c]e^2$ to be unity and, in effect ignore all physical constants appearing in our expressions. Note that we have in reserve the ability to set yet one other constant to unity in the future, if necessary.

6 References

[1] P. Hohenberg and W. Kohn Phys. Rev. **136**, B864 (1964).
[2] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).