

The behavior of a material can be related to the types of bonding between the atoms, whether it be metallic, covalent, ionic, etc. That bonding represents the distribution of electrons around the nuclei. Covalent bonds have a localized electronic distribution between atoms and are generally strong and directional. Materials with strongly covalent bonds include important semiconductors, such as silicon, gallium, and diamond. Metallic systems, in contrast, may have a degree of directionality to their bonding, but the dominant feature is a delocalized sea of electrons. Ionic bonds are dominated by the strong electrostatic interactions between the ions. Fundamentally, the properties of each material start with its bonding.

A fundamental description of bonding requires a calculation of the electronic distributions. The class of methods that yield such information are called *electronic structure methods*. In this chapter, we shall briefly review the basics of these methods, pointing out their inherent approximations. There are numerous books devoted to the fundamental theories behind these methods – embodied in quantum mechanics – as well as many texts devoted to electronic structure methods themselves [167, 219, 251, 254]. We can at best give a brief guide to this topic needed for discussions later in the text and as well as for a basis for understanding and evaluating this fascinating field.

Not so many years ago, practitioners of electronic structure calculations typically used home-grown computer codes, which often required heroic efforts on the parts of the programmers. These days, numerous codes are available, ranging from being free to being quite expensive and thus almost no one writes their own codes any more. The good news is that electronic structure calculations are thus now widely available and accessible to most researchers. The negative side is that there are many chances to do poor-quality calculations. Please be warned that the discussion in this text will not adequately prepare you to avoid all the pitfalls associated with these calculations.

This section may be a bit challenging for those without much experience in quantum mechanics. We have provided an all-too-brief introduction in Appendix F. The interested reader is urged to go beyond that simple description and read some of the texts listed at the end of this chapter. A nice tutorial on electronic structure calculations, with greater sophistication than this chapter, is available from [240].

4.1 QUANTUM MECHANICS OF MULTIELECTRON SYSTEMS

The basis for all quantum mechanical calculations of the electronic structure of a material is the Schrödinger equation,

$$\mathcal{H}\Psi = E\Psi, \quad (4.1)$$

where \mathcal{H} is the Hamiltonian operator, E is the energy, and Ψ the wave function.¹ A key point of quantum mechanics is that \mathcal{H} is an operator and thus behaves quite differently than an equivalent function in classical systems. Some details about operators are given in Appendix F.3.

The Schrödinger equation describes the energy of the electrons and nuclei in a material. Electrons interact with the positively charged atomic nuclei through an electrostatic potential. From the point of view of the electrons, the nuclei are fixed,² thus the electron interaction with a nucleus (α) can be considered as an external potential that takes the form $v = -Z_\alpha/r_{i\alpha}$, where Z_α is the nuclear charge and $r_{i\alpha}$ is the distance from the nucleus α to the i th electron. If a system has N electrons and M nuclei, the Hamiltonian is

$$\mathcal{H} = \sum_{i=1}^N \left(-\frac{1}{2} \nabla_i^2 + v_{ext}(\mathbf{r}_i) \right) + \sum_i \sum_{j>i} \frac{1}{r_{ij}} \quad (4.2)$$

where the electrostatic (Coulomb) potential acting on electron i from the M nuclei is a sum over all M nuclei

$$v(\mathbf{r}_i) = - \sum_{\alpha=1}^M \frac{Z_\alpha}{r_{i\alpha}}. \quad (4.3)$$

These equations are in *atomic units*, in which the fundamental constants m , e , \hbar , etc. are all equal to 1. The unit of energy is the hartree, the unit of length is the bohr, etc., as discussed in more detail in Appendix F.8.

The wave function Ψ depends on the position of the electrons, \mathbf{r} . We will be more specific later, but for now let us just write the wave function as a function of the position of the N electrons as $\Psi = \Psi(\mathbf{r}^N)$, where \mathbf{r}^N is shorthand notation for the quantity $\{\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N\}$. The electron *density* of the system, ρ , is the number of electrons per unit volume,

$$\rho(\mathbf{r}_1) = N \int \dots \int |\Psi(\mathbf{r}^N)|^2 d\mathbf{r}_2 \dots d\mathbf{r}_N, \quad (4.4)$$

where we integrate over all electron coordinates but one. The normalization of the wave functions is $\int \rho(\mathbf{r}_1) d\mathbf{r}_1 = N$.

The Schrödinger equation in Eq. (4.2) is the fundamental equation describing the electronic structure of materials. If it could be solved exactly, the solution would be the wave function and

¹ More details are given in Appendix F and Eq. (F.6).

² Essentially all calculations are done based on the Born-Oppenheimer approximation, in which it is assumed that since electrons are much lighter and smaller than the atomic nucleus, they move fast enough that they can respond "instantaneously" to changes in the atomic positions. Thus, the nuclear coordinates can be taken as fixed when determining the electron distributions and energies.

energy, giving a complete description of the electronic properties. However, we cannot solve these equations exactly except for some simple problems.³ The main objective of this chapter is to introduce the approximations used in modern electronic structure methods.

4.1.1 Classes of methods

There are classes of methods that assume an approximate wave function and find variational solutions to Eq. (4.2). These methods go by many names, depending on the approximations used for the wave functions. For example, in the *Hartree* approximation, the total wave function consists of single particle functions for each electron, which leads to an inaccurate, but easy to solve, set of equations. The electronic wave functions suffer from not being antisymmetric with respect to the exchange of electrons, as discussed in Appendix F.6.2.

In *Hartree-Fock* calculations, the wave functions are constructed so that they are antisymmetric with respect to the exchange of electrons and the results from these calculations are much superior to that with the simpler Fock method. Hartree-Fock theory neglects, however, the *correlation* energy, which accounts for the energy associated with the motions of the electrons being correlated so that they stay apart from each other (Appendix F.6).

Hartree-Fock and other, similar, methods were the mainstay of electronic structure calculations (especially in quantum chemistry) for decades, but have largely been supplanted by an approach based on a *density functional theory* (DFT) formulation of quantum mechanics. In DFT, the energy is written as a function of the electronic density, which is a function of the position, thus the energy is a *functional* of the density.⁴ Given its ubiquitous use, DFT will be the focus of the rest of this chapter. Those interested in more traditional methods are urged to peruse the book by Kaxiras listed in the Suggested readings at the end of the chapter.

4.2 EARLY DENSITY FUNCTIONAL THEORIES

The earliest attempts to use the electronic density as a fundamental parameter in an electronic-structure calculation were published in 1927 independently by Thomas [309] and Fermi [99] (note that Schrödinger's equation was only published in 1926). They were each searching for a simple way to solve for the electronic structure of atoms. By analogy with classical mechanics, they assumed that there were three components to the energy: the attraction between the electrons and the nucleus of the atom, the kinetic energy of the electrons, and the repulsion between the electrons. They then assumed they could write the energy not in terms of wave functions, but rather in terms of functions of the electronic density of the system, $\rho(\mathbf{r})$.

Electrons interact with the positively charged atomic nuclei through an electrostatic potential, which for an atom takes the form $v_{ext} = -Z/r$, where Z is the nuclear charge and r is the distance from the nucleus to the electron. The interaction energy of all the electrons with the

³ A few examples are discussed in Appendix F.5.

⁴ See Appendix C.6 for a description of the properties of functionals.

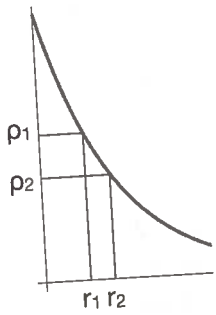


Figure 4.1 Local-density approximation. Although the electronic density is rapidly varying, the energy expression at each density is based on a uniform electron gas with that electronic density.

nucleus is found by evaluating the integral of the electronic density as a function of r times the interaction of an electron a distance r away from the nuclei, i.e.,

$$\int \rho(\mathbf{r})v_{ext}(\mathbf{r})d\mathbf{r}. \quad (4.5)$$

In quantum mechanics, the kinetic energy is calculated with the $-(1/2)\nabla^2$ term in the Hamiltonian in Eq. (4.2). In the Thomas-Fermi model, the kinetic energy density is approximated by the expression for the energy of a uniform electron gas with density ρ , as described in Appendix F.5.1. The total kinetic energy of the electron gas, T_{eg} , is the integral over space of the kinetic energy density and in atomic units takes the form (from Eq. (F.24))

$$T_{eg}[\rho] = C_F \int \rho^{5/3}(\mathbf{r})d\mathbf{r}, \quad (4.6)$$

where $C_F = (3/10)(3\pi^2)^{2/3}$.

There is an aspect to Eq. (4.6) that deserves some discussion. In Figure 4.1 we show a schematic view of the electronic density near the nucleus of an atom. At a distance r_1 in that figure, the density is ρ_1 , and at r_2 , it is ρ_2 . In Eq. (4.6), the kinetic energy density at the two distances will be proportional to $\rho_1^{2/3}$ and $\rho_2^{2/3}$, respectively. However, the expression for the kinetic energy density was derived assuming a gas of electrons with uniform density everywhere (see Appendix F.5.1). Applying the electron-gas expression at each point regardless of how the density is actually varying is an approximation. However, it is an approximation that is essential to many electronic structure calculations. Since the density functionals are applied at each point and depend only on the values of the electronic density at that point, this approach is referred to as a *local-density approximation* (LDA).

The final term in the Thomas-Fermi (TF) model is the classical electrostatic (Coulomb) interaction between the electrons, a term that occurs so regularly it has its own symbol, J .⁵ J is given by an integral over the electronic distribution

$$J[\rho] = \frac{1}{2} \iint \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_2 - \mathbf{r}_1|} d\mathbf{r}_1 d\mathbf{r}_2. \quad (4.7)$$

The total energy in the TF model is

$$E_{TF}[\rho] = T_{eg}[\rho] + J[\rho] - Z \int \frac{\rho(\mathbf{r})}{r}. \quad (4.8)$$

⁵ Described in Appendix E.2.

E_{TF} is a *functional* of the electron density ρ ,⁶ that is, it is a function of the function ρ .⁷ The Thomas-Fermi model is thus a simple example of a *density functional theory* (DFT).

In quantum mechanics, there are corrections to the Coulomb term (J in Eq. (4.8)) that arise from the quantum nature of electrons. One of those corrections comes from a quantum effect called *exchange*, which arises from the fact that two electrons of the same spin can never be found at the same place. Thus, electrons of the same spin avoid each other and the electron-electron repulsion energy is less for electrons with the same spin than for electrons with the opposite spin. This difference in energy between parallel-spin and anti-parallel-spin electrons is the exchange energy. The origins of the exchange energy are discussed in Appendix F.6.3, in which we show that the exchange energy for a uniform electron gas is given by an integration of the exchange energy density from Eq. (F.45) over space,

$$E_x[\rho] = -C_x \int \rho^{4/3}(\mathbf{r})d\mathbf{r}, \quad (4.9)$$

where $C_x = (3/4)(3/\pi)^{1/3}$.

In 1930, Dirac modified the Thomas-Fermi model Eq. (4.8) to include the exchange energy [92]. Called the Thomas-Fermi-Dirac (TFD) model, the total energy functional is

$$E_{TFD}[\rho] = T_{eg}[\rho] + E_x[\rho] + J[\rho] - Z \int \frac{\rho(\mathbf{r})}{r} d\mathbf{r}. \quad (4.10)$$

The goal in both the Thomas-Fermi and Thomas-Fermi-Dirac models was to determine the ground state (lowest) energy and the ground state electronic density, which was accomplished by minimizing the energy with respect to the density. The idea that the ground state could be determined through such a minimization was an assumption, not proved until the Hohenberg-Kohn theorem described in the next section. The minimum can be found using the method of Lagrange multipliers with the constraints of fixed total number of electrons (the integral of the density), the requirement that $\rho(\mathbf{r}) \rightarrow 0$ as $\mathbf{r} \rightarrow 0$ (to avoid a singularity in the Coulomb energy), and that $\rho(\mathbf{r}) \rightarrow 0$ as $\mathbf{r} \rightarrow \infty$ (finite-sized atoms).⁸

⁶ Some properties of functionals are described in Appendix C.6.

⁷ Functionals have some interesting properties, which are reviewed in Appendix C.6. We will see functionals again in the discussion of the phase-field model in a later chapter.

⁸ Consider the solution to the TFD model. We can include the constraint that $\int \rho(\mathbf{r}_1)d\mathbf{r}_1 = N$ through the use of Lagrange multipliers [10], which leads to

$$\delta \left\{ E_{TFD}[\rho] - \mu_{TFD} \left(\int \rho(\mathbf{r}_1)d\mathbf{r}_1 - N \right) \right\} = 0, \quad (4.11)$$

where δ is a functional differential (see Appendix C.6 for a discussion of the calculus of functionals). Taking the functional derivatives (e.g., $\delta E/\delta\rho$), we have

$$\mu_{TFD} = \frac{\delta E[\rho]}{\delta\rho} = \frac{5}{3}C_F\rho^{2/3}(\mathbf{r}) - \frac{4}{3}C_x\rho^{1/3}(\mathbf{r}) - \phi(\mathbf{r}) = 0, \quad (4.12)$$

where the electrostatic potential is

$$\phi(\mathbf{r}) = \frac{Z}{r} - \int \frac{\rho(\mathbf{r}_1)}{|\mathbf{r} - \mathbf{r}_1|} d\mathbf{r}_1. \quad (4.13)$$

Eq. (4.12) can be solved in a number of ways. For example, a functional form could be introduced that includes the constraints that $\rho(\mathbf{r}) \rightarrow 0$ as $\mathbf{r} \rightarrow 0$ and $\rho(\mathbf{r}) \rightarrow 0$ as $\mathbf{r} \rightarrow \infty$ or strictly numerical solutions could be used.

The TF and TFD models are appealingly simple, being based on functionals from a uniform electron gas and used with a local density approximation. They are, however, based on a crude description of the kinetic and exchange energies, they neglect the correlation energy described in Appendix F.6, their electronic densities are not based on realistic wave functions, and so on. Given these approximations, the question then is: how accurate are these methods?

There are actually two questions we could consider: (1) how well do the results from the TF and TFD calculations match similar, but more accurate, calculations? and (2) how well do the results match the real energies of atoms and molecules? This type of question will face us throughout this text. Should the results calculated from models be compared to reality or other calculations? Often, it will be the latter. For example, the TF and TFD models are exactly that, models. They neglect important physics. Thus, they cannot be expected to match experiments or accurate calculations on real systems. So question (2) is not very relevant in this case. A more informative approach would be to compare these model calculations with accurate calculations using other methods that most closely match these, which will give us a sense of how well these models work. In the case of the TF and TFD models, they can best be compared to accurate calculations using methods such as the Hartree-Fock approach mentioned in Section 4.1.1. The Hartree-Fock method solves the Schrödinger equation in Eq. (4.2) very accurately, using antisymmetric wave functions. The Hartree-Fock method does not, however, include the correlation energy. Thus, comparison of TF and TFD to Hartree-Fock results is a direct measure of how well these simple models capture the physics of multielectron systems without correlation. It can be shown analytically that the Thomas-Fermi energy for neutral, closed-shell, atoms is [251]

$$E_{TF} = -0.7687 Z^{7/3}, \quad (4.14)$$

where Z is the charge on the nucleus (and equals the number of electrons). In Table 4.1 we compare the Thomas-Fermi energy to accurate calculations with the Hartree-Fock method, where we plot $-E_{HF}/Z^{7/3}$. If the Thomas-Fermi result were correct, then the Hartree-Fock result should take the value 0.7687 from Eq. (4.14). From the table, it is clear that the Thomas-Fermi model yields energies with considerable error, overestimating the binding energy of the electrons, with relative errors ranging from about 35% for He to 13% for Rn.

While adding the exchange correction was supposed to improve the TF results, they actually make predictions of TFD worse than TF, which we can see directly from Eq. (4.10). The Thomas-Fermi model overestimates the binding energies of the atoms. Adding an exchange term that is negative will lower the energy even more, increasing the error.

The TF and TFD models do poorly when compared to accurate results without correlation energy. A number of suggested improvements have been proposed to these models, including the use of gradient terms in the kinetic energy [251]. While the results are somewhat better, they are still not adequate. Thus, there is something fundamentally incorrect about the models. The major error arises from the treatment of the kinetic energy, which, in the TF and TFD models, ignores the shell structure of the electrons [219].

Despite their failings, we discuss the TF and TFD models because they represent the first of a class of methods that are based on what is called *density functional theory* (DFT) for calculating electronic structure. There has been a long history of these and similar methods, all based on an

Table 4.1 Comparison of the energy of the Thomas-Fermi model with calculated energies E_{HF} for rare-gas atoms [70]. The calculated energies were determined with the Hartree-Fock method, so do not contain the correlation energy. From [251]

| Atom | Z | $(-E_{HF}/Z^{7/3})$ |
|------|-----|---------------------|
| He | 2 | 0.5678 |
| Ne | 10 | 0.5967 |
| Ar | 18 | 0.6204 |
| Kr | 36 | 0.6431 |
| Xe | 54 | 0.6562 |
| Rn | 86 | 0.6698 |

assumption that one could find a true functional of the density that would describe the energy of a system of electrons and nuclei. This assumption was a leap of faith. Indeed, these methods were often mocked as not being based on good theory. It was not until the Hohenberg-Kohn theorem in the early 1960s that one knew for certain that such a density functional actually existed.

4.3 THE HOHENBERG-KOHN THEOREM

In a remarkable theorem, Hohenberg and Kohn [149] (HK) showed that the total energy, E , of a system of electrons in an external potential (in this case the Coulomb potential from the nuclei in a solid) is given *exactly* as a functional of the electronic density ρ . They showed further that the density that minimizes $E[\rho]$ is the ground-state electronic density and that other ground-state properties are also functionals of the ground-state density. Through this theorem, they proved that TF, TFD, and other such approaches had a theoretical justification – the Hohenberg-Kohn (HK) theorem made density functional theory well-founded.

The HK theorem, for all its importance, begs an important question. While the HK theorem states that a functional $E[\rho]$ exists, it fails to tell us what that functional is or how to find it. What the HK theorem does do, however, is to say that it is worth looking for such a functional.

Models such as TF, TFD, etc. were attempts at finding the correct functional. They were not very good, perhaps, but they were a step in the right direction. In the next section, we will discuss a more accurate approach that is the basis for most calculations today.

4.4 KOHN-SHAM METHOD

In this section we introduce the Kohn-Sham (KS) approach to solving the quantum mechanics of multielectron systems [173]. The use of this model, coupled with advances in treating the

correlation energy, solves many of the inadequacies of the TF/TFD models and is the basis of most DFT calculations done today. Much of this discussion is based on that in [251]. A more recent review [240] is also quite useful. For more advanced readers, other recent books may be of help [102, 219].

The functional $E[\rho]$ represents the total electronic energy as a functional of the electron density ρ , and is a sum of the contributions from the external potential and the electronic energies,

$$E[\rho] = F[\rho] + \int v_{\text{ext}}(\mathbf{r})\rho(\mathbf{r})d\mathbf{r}, \quad (4.15)$$

where

$$F[\rho] = T[\rho] + V_{ee}[\rho]. \quad (4.16)$$

$F[\rho]$ is a sum of the kinetic energy $T[\rho]$ and the interaction energy between the electrons $V_{ee}[\rho]$. We write $V_{ee}[\rho]$ as the classical Coulomb integral (Eq. (E.10)) plus a correction term

$$V_{ee}[\rho] = J[\rho] + (V_{ee}[\rho] - J[\rho]). \quad (4.17)$$

The focus is now on finding approximations for the kinetic energy T and the part of the electron-electron interaction that is not described by the classical Coulomb energy J , i.e., $V_{ee}[\rho] - J[\rho]$.

As noted earlier, a major source of error in the TF and TFD models arises from the kinetic energy not reflecting the discontinuities inherent in the shell structure of the electrons. To avoid these difficulties, KS assumed that electron density of a system with N electrons could be written as a sum of *one-electron orbitals* ψ_i as

$$\rho(\mathbf{r}) = \sum_{i=1}^N |\psi_i(\mathbf{r})|^2. \quad (4.18)$$

The significance of writing the density in one-electron orbitals is that (1) finding solutions to the Schrödinger equation is greatly simplified and that (2) the discontinuities in the shells are natural outcomes of the solutions. As shown below, the multielectron problem is reduced to finding a set of solutions to one-electron problems.

Kohn and Sham defined the kinetic energy functional to be

$$T_{KS}[\rho] = \sum_{i=1}^N \langle \psi_i | -\frac{1}{2}\nabla_i^2 | \psi_i \rangle, \quad (4.19)$$

where we use the notation from Eq. (4.26). The point is that KS assumes a simple orbital wave function picture for the density and then approximates the kinetic energy with the exact kinetic-energy function operating on those approximate wave functions. The electronic part of the density functional in Eq. (4.16) is then given by

$$F[\rho] = T_{KS}[\rho] + J[\rho] + E_{xc}[\rho], \quad (4.20)$$

where

$$E_{xc}[\rho] = T[\rho] - T_{KS}[\rho] + V_{ee}[\rho] - J[\rho]. \quad (4.21)$$

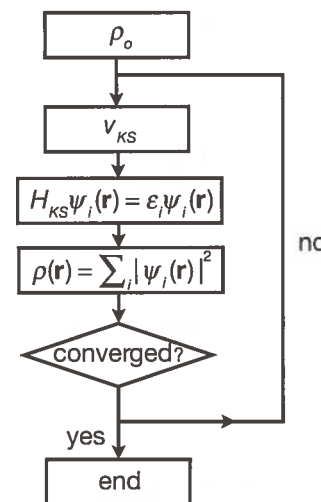


Figure 4.2 Steps in the solution to the Kohn-Sham method. Adapted from [240].

The term E_{xc} in Eq. (4.21) is called the exchange-correlation energy. It includes all the corrections between the sum of the kinetic and Coulomb energies and the correct answer. The hope is that E_{xc} , as a correction term, is small relative to the rest of the terms in Eq. (4.20). The task is then to develop a good form for E_{xc} , as discussed in the next section.

The KS method uses an iterative solution, shown schematically in Figure 4.2. One initially guesses at a set of wave functions ψ_i^0 , from which is constructed an initial electron density ρ_0 from Eq. (4.18). The effective KS potential based on this density is found by taking the functional derivative of $E[\rho]$ yielding⁹

$$v_{KS}(\mathbf{r}) = v_{\text{ext}}(\mathbf{r}) + \int \frac{\rho(\mathbf{r}_1)}{|\mathbf{r} - \mathbf{r}_1|} d\mathbf{r}_1 + v_{xc}(\mathbf{r}), \quad (4.22)$$

where $v_{xc} = \delta E_{xc} / \delta \rho(\mathbf{r})$ and E_{xc} is the exchange-correlation function of choice.

Since v_{KS} is defined by the electron density from the previous step in the iteration cycle, the Hamiltonian includes no direct interactions between electrons – it describes each electron as moving in an external field based on a fixed electron distribution, $\rho(\mathbf{r})$. The Hamiltonian is

$$\mathcal{H}_{KS} = -\frac{1}{2}\nabla^2 + v_{KS}(\mathbf{r}), \quad (4.23)$$

from which we solve

$$\mathcal{H}_{KS} \psi_i(\mathbf{r}) = \epsilon_i \psi_i(\mathbf{r}). \quad (4.24)$$

Solving Eq. (4.24) yields a new set of N orbital wave functions ψ (and the 1-electron orbital energies ϵ), from which a new ρ is found, and then a new v_{KS} . Equation (4.24) is solved again, and the process is repeated until a self-consistent ρ is found, by which we mean a ρ that does not vary more than a prescribed amount from one iteration to the next.

⁹ Functional derivatives are described in Appendix C.6.

Typically, one expands the wave function ψ in a set of basis functions,

$$\psi = \sum_j c_j \phi_j. \quad (4.25)$$

To solve Eq. (4.24), we form the Hamiltonian matrix and then diagonalize it to find the eigenvectors and eigenvalues.¹⁰ The Hamiltonian matrix elements are

$$H_{ij} = \int \phi_i^*(\mathbf{r}) \left\{ -\frac{1}{2} \nabla^2 + v_{KS}(\mathbf{r}) \right\} \phi_j(\mathbf{r}) d\mathbf{r}. \quad (4.26)$$

The energy of the KS method is not the sum of the "orbital" energies ϵ_i from Eq. (4.24) – there are additional terms that must be added to arrive at the correct total energy. The total energy is [240]¹¹

$$E = \sum_i^{n_{occ}} \epsilon_i - \int \left[\frac{1}{2} \int \frac{\rho(\mathbf{r}_1)}{|\mathbf{r} - \mathbf{r}_1|} d\mathbf{r}_1 + v_{xc}(\mathbf{r}) \right] \rho(\mathbf{r}) d\mathbf{r} + E_{xc}. \quad (4.27)$$

It is important to emphasize that the iterative method in Figure 4.2 converges to the minimum-energy state for choice of the basis set employed in the calculation.¹² A poor basis set will lead to poor calculations of the ground-state of the system. Indeed, one test of a basis set is to compare the ground-state energy with that calculated with other basis sets, with the descriptions that yield the lowest energy usually being preferred.

4.5 THE EXCHANGE-CORRELATION FUNCTIONAL

In Eq. (4.21) is defined a term called the *exchange-correlation* functional, $E_{xc}[\rho]$. This term includes all the corrections to the approximate Kohn-Sham Hamiltonian. There are two main types of exchange-correlation functionals in use, those based on the local-density approximation (LDA) and methods that include corrections to LDA, usually in the form of gradients of the density.

The TFD model is an example of a local-density approximation for the exchange function E_x , given in Eq. (4.9). Common approximations for E_{xc} start with E_x and add a correlation term,

$$E_{xc} = E_x + E_c. \quad (4.28)$$

There have been a number of models used for E_c . A common one is a simple functional of ρ , much as is E_x . Exacting quantum Monte Carlo calculations¹³ were performed for a uniform

¹⁰ Working with basis sets is described in Appendix F.7.

¹¹ Note that $\int E_{xc} d\mathbf{r} \neq \int v_{xc} \rho(\mathbf{r}) d\mathbf{r}$.

¹² That the minimum energy corresponds to the ground state is a consequence of the Hohenberg-Kohn theorem and the structure of the Kohn-Sham equations.

¹³ The quantum Monte Carlo method is an approach to solving the Schrödinger equation very accurately as described in [60].

electron gas and the correlation energy was calculated as the total energy minus the kinetic, coulomb, and exchange energies [59]. Accurate expressions for the energy were obtained at high and low electron densities and a variety of schemes were proposed to parameterize these results [256, 257]. The advantage to these methods is that E_{xc} remains a simple functional of ρ and so the same types of manipulations can be followed to derive $v_{xc} = \delta E_{xc}[\rho(\mathbf{r})]/\delta \rho(\mathbf{r})$. A disadvantage is that in these schemes, E_{xc} is still calculated in the local-density approximation.

The most common corrections to E_{xc} are still local theories, but are based on not just the local value of the electron density but also on the local value of the gradient of the electron density. These methods are typically designated as *generalized gradient approximations* (GGA). In these methods, the exchange-correlation functional is assumed to have the form (for example in [191, 258])

$$E_{xc} = \int \rho(\mathbf{r}) \epsilon_{xc}(\rho, \nabla \rho) d\mathbf{r}, \quad (4.29)$$

where $\nabla \rho(\mathbf{r})$ is the gradient of $\rho(\mathbf{r})$. The derivation of the expression for v_{xc} is not difficult, as the functional derivatives are straightforward.

Various other approximations for E_{xc} have been proposed,¹⁴ but LDA and GGA are the most commonly used today for solids. As discussed below, both LDA and GGA functionals yield band gaps that are too small, so there is a focus on finding new functionals. Hybrid approaches, in which there is an interpolation between the Hartree-Fock approximation and the GGA [191], are increasingly being used. While initially developed for chemical applications, they have also been implemented for solids, in which they yield greatly improved descriptions of the electronic structure.

4.6 WAVE FUNCTIONS

The first step in a calculation is to choose the wave function. As discussed in Eq. (4.25), one typically writes the wave function as a sum over a set of functions, referred to as the basis set. What goes into the choice of basis set depends on the answers to a couple of questions:

- Can results with sufficient accuracy be found with the choice of basis set?
- What is the computational cost for achieving converged results with the desired accuracy?

Not surprisingly, the choice of wave function depends largely on the system of interest.

For solids, the wave functions should reflect the periodic symmetry of the system – they must satisfy Bloch's theorem [15] as described in Appendix F.9¹⁵

$$\phi_{\mathbf{k}}(\mathbf{r} + \mathbf{R}) = e^{i\mathbf{k} \cdot \mathbf{R}} \phi_{\mathbf{k}}(\mathbf{r}), \quad (4.30)$$

¹⁴ Including meta-GGA methods which include the Laplacian ($\nabla^2 \rho$) in the functional [238].

¹⁵ For more discussion of periodic systems, please see Section 3.4, in which periodic boundary conditions are introduced.

where \mathbf{R} is a direct lattice vector. A general set of functions that satisfy this condition are *plane waves*

$$\phi_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} \sum_{\mathbf{G}} c_{\mathbf{G}} e^{i\mathbf{G}\cdot\mathbf{r}}, \quad (4.31)$$

where \mathbf{G} is a reciprocal lattice vector. The energy of a specific plane wave is proportional to k^2 , as seen in the solution to the particle in a box in Appendix F.5.1. For application in electronic structure calculations, a set of wave functions with a choice of k values is chosen and the coefficients in the wave function are determined as described in Figure 4.2.

Plane wave basis sets are common solutions for free electrons with boundary conditions that restrict the periodicity of the wave function¹⁶ and offer many advantages in density-functional calculations for solids. The integrals involving plane waves are easier to evaluate and code than more complicated basis functions. They form a complete set and, by including sufficient numbers of them, yield convergence to an arbitrary accuracy. In practice, a finite number of plane-wave functions is used, with the convergence of the properties governed largely by a single parameter, the maximum energy plane wave used in the basis set.¹⁷ Calculations based on a plane-wave basis generally converge smoothly to the minimum-energy solution. Since all functions in a plane-wave basis are mutually orthogonal, their use avoids a problem called “basis-set superposition error” that arises when non-plane-wave basis sets are used, as discussed below. Many more details are given in [240].

Plane waves can also be used for finite systems, such as atoms or molecules, by means of a “supercell” approach, in which the finite system is placed in a unit cell of a fictitious crystal, which is then made large enough such that the atoms and molecules in the central unit cell do not interact with their replicas in other cells. Supercells are used to model not just finite systems, but any system for which three-dimensional periodicity is broken. Examples include surfaces or interfaces (with two-dimensional periodicity), nanowires (one-dimensional periodicity) and quantum dots (with no periodicity).

While plane wave basis sets are common, they are not necessarily the best choice for all problems. For example, when describing solids consisting of molecules or systems with strongly directional bonding, it is common to use a basis of a finite number of atomic orbitals rather than a set of plane waves. These orbitals are centered at the atomic nuclei, making this choice of basis set similar to (but not actually the same as) the Linear Combination of Atomic Orbitals (LCAO) approximation that has been in use for many years to describe molecular electronic structure. While hydrogenic-like orbitals conform more closely to actual electronic distributions, approximating these with a linear combination of Gaussian functions offers many advantages owing to the ease of calculating the various integrals needed to evaluate the Hamiltonian matrix elements. In Figure 4.3 we compare a hydrogenic-orbital (called a Slater function) to an equivalent Gaussian function. Note the significant differences both at short and long range. The hope is that even with the Gaussian functions, many fewer basis set functions

¹⁶ See, for example, the particle in a box in Appendix F.5.1.

¹⁷ To avoid problems associated with the rapidly varying electronic density near the nuclei of the atoms, approximate functions, called pseudopotentials, are used to approximate the core regions, as discussed in the next section.

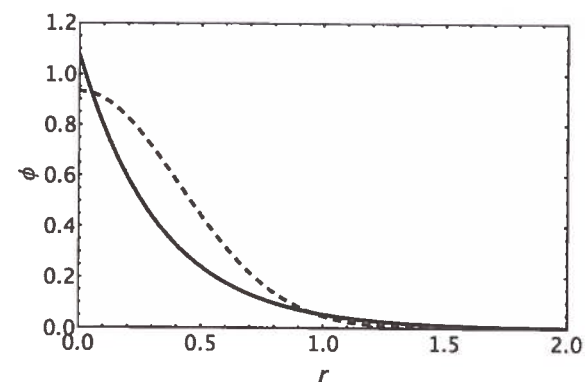


Figure 4.3 Comparison of Slater and Gaussian orbitals. Solid curve is a normalized atomic-orbital-type (Slater) orbital of the form e^{-ar} and dashed curve is a normalized Gaussian orbital of the form e^{-ar^2} , with $\alpha = 2$.

will be needed than in plane wave calculations because each function corresponds more closely to the actual form of the electronic distribution. Thus, non-plane-wave basis approaches tend to be computationally fast. Another advantage of the use of non-plane-wave basis functions is that they do not need to be periodic, which enables the calculation of properties of aperiodic or semiperiodic supercells.

The actual choice of functions that make up a basis set depends on a number of factors, with the final choice usually being governed by a balance between accuracy and computational time. The danger is that since one does not use a complete set of functions, the solutions found in the minimization of the energy may be incorrect owing to having excluded the actual ground state from the solution by the choice of basis set. It is also sometimes required to add “ghost” atoms, which are atomic-like orbitals that are not centered on an atom but instead are added in an ad hoc fashion to describe a region of space near atoms. For example you might add extra orbitals near point defects, surfaces, or in the covalent bond-charge. Details of these methods are beyond the scope of this text.

Calculations based on a finite basis set often suffer from “basis set superposition error”, which arises from the overlap of their basis functions as two atoms come close to each other. When this occurs, the wave function of an atom begins to include contributions from the wave function of different atoms, leading to an overall error in the calculation. There are a couple of ways to eliminate or reduce this error, the details of which we leave to more advanced texts.

Another method that does not rely on plane waves is based on real space grids. At each point on the grid the wave function is just a numerical value and the Kohn-Sham Hamiltonian is solved numerically. This approach can be very fast and is easily extendable to non-periodic systems.

4.7 PSEUDOPOTENTIALS

It is common to think about the electronic structure of atoms, whether in a solid or not, by dividing the electrons into two groups, the outer, valence, electrons, and the electrons at the

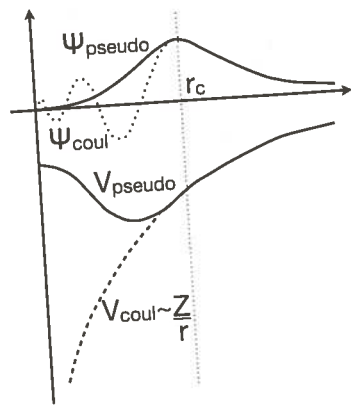


Figure 4.4 Schematic view of a pseudopotential and a pseudo wave function, as described in the text. The bottom part of the plot shows the potential, with the term V_{coul} representing the bare Coulomb potential from the core electrons and V_{pseudo} the pseudopotential. In the top part of the figure we show the wave functions, with Ψ_{coul} showing the oscillations in the wave function that arise because the wave functions of the valence electrons must be orthogonal to those of the core electrons. By taking the core electrons out of the problem, the oscillations can be eliminated, leading to Ψ_{pseudo} , the smooth wave function that arises from the pseudopotential. r_c is the core radius.

inner core. The inner shell electrons are very tightly bound to the nucleus and play little to no role in the chemical binding between atoms. In many ways, the inner shell electrons plus the nucleus can be considered as an essentially inert core, with a charge that is the normal charge on the nucleus minus the charge associated with the inner core electrons. The advantage to this point of view is that the inner core electrons can then be considered as frozen and only the outer, valence, electrons need be included in the calculation of the electronic structure. Of course, since the inner electrons are not point charges, they cannot be ignored completely.

The effects of the inner electrons are modeled by the use of *pseudopotentials*, in which each nucleus plus its core electrons is treated as a frozen core that does not change in response to changes in its environment. The interactions of the valence electrons with this core are described by a potential function, called the pseudopotential, that may be constructed from first-principles calculations to reproduce the behavior of the nucleus and core electrons exactly. The accuracy of pseudopotential calculations has been justified by comparison with calculations that include all of the electrons. Pseudopotential methods are much faster than all-electron calculations because the tightly-bound core electrons do not enter directly into the problem [290].

The simplest are the so-called empirical pseudopotentials, in which a functional form is assumed and parameters are chosen to match experiment. *Ab initio* pseudopotentials are determined from accurate calculations of the free atom Schrödinger equation. The empirical pseudopotentials are not commonly used today, with first-principles pseudopotentials, based on local solutions to the wave functions near the atomic nuclei, being the preferred approach. A detailed description of pseudopotentials can be found in [240].

In Figure 4.4 we show a schematic view of how a pseudopotential acts in the system. The bottom part of the plot shows the potential, with the term V_{coul} representing the bare Coulomb potential from the core electrons and V_{pseudo} the pseudopotential, in which the Coulomb potential is offset by the core electrons. In the top part of the figure we show the wave functions, with Ψ_{coul} showing the oscillations in the wave function owing to the steep Coulomb potential and Ψ_{pseudo} showing the smooth wave function that arises from the pseudopotential. r_c is the core radius.

Table 4.2 Comparison of some bulk properties of bulk Si obtained with LDA and GGA [240]. a is the lattice constant, B the bulk modulus, and E_c the cohesive energy. Experimental results are from [192]

| | LDA | GGA | Experiment |
|-----------------|-------|-------|------------|
| a (Å) | 5.378 | 5.463 | 5.429 |
| B (Mbar) | 0.965 | 0.882 | 0.978 |
| E_c (eV/atom) | 6.00 | 5.42 | 4.63 |

Table 4.3 Comparison of some bulk properties of bulk Cu obtained with LDA and GGA [240]. a is the lattice constant, B the bulk modulus, and E_c the cohesive energy. Experimental results are from [277]

| | LDA | GGA | Experiment |
|-----------------|-------|-------|------------|
| a (Å) | 3.571 | 3.682 | 3.61 |
| B (Mbar) | 0.902 | 0.672 | 1.420 |
| E_c (eV/atom) | 4.54 | 3.58 | 3.50 |

4.8 USE OF DENSITY FUNCTIONAL THEORY

The density functional methods described in this chapter can reproduce the structural properties of materials with reasonable accuracy. Typically, LDA calculations underestimate equilibrium lattice parameters by about 1-2%, while use of the GGA leads to larger bond lengths than LDA, with predicted values that can be somewhat larger than experiment. Similarly, LDA predicts a cohesive energy larger than the experimental value, with the GGA correcting the predicted value.

For example, in Table 4.2 we compare calculated values (LDA and GGA) with experimental values of structural properties for bulk Si, a covalently bonded material. The LDA results are too bound, with a lattice parameter about 1% low, and the GGA results are a bit too large, though closer to experiment. The LDA and GGA results for the cohesive energy are both too large. In Table 4.3 we show similar results for a metallic system, bulk Cu. Again, the LDA results for the lattice parameter are too small (about 1%), while the GGA results are too large. Both methods give cohesive energies that are too large, though GGA yields results that are closer to experiment. The bulk modulus is not well calculated with either method for either system, being too small, though LDA does somewhat better than GGA.

While the trends in the differences in calculations based on LDA and GGA in Table 4.3 seem to generally hold, they are not always followed. Indeed, there is no theory that explains the

difference between the results of the two methods. It is somewhat surprising that GGA-based calculations seem no more accurate than LDA-based ones. One might expect that including any additional information, such as the charge gradient, would result in a more accurate calculation.

The Hohenberg-Kohn theorem, and the Kohn-Sham method based on it, is only valid for the ground state energy. Thus it is not surprising that DFT does not do well in calculating the energy of excited states. For example, band gaps are predicted to be too small by approximately a factor of 2 in most cases. In Si, for example, the experimental gap is 1.17 eV and that calculated with LDA is 0.45 eV and the value with GGA is 0.61 eV [240].

One thing DFT does very well is to determine the electron density, which can lead to new insights into how materials are structured and behave. As just one example, in a study using DFT to study the potential of certain Bi compounds (BiAlO_3 and BiGaO_3) as piezoelectric materials [18], it was found that the piezoelectric response is a result of the stereochemical activity of the Bi lone pairs, which causes large displacements of the Bi ions from their positions in a centrosymmetric phase. This insight was gained from an analysis of plots of the electron localization.

Other limitations of DFT center on its computational complexity, which limits the number of atoms that can be included in a calculation. While that number changes with time as computers are made faster and algorithms are improved, suffice it to say that there are many problems for which these calculations are not well suited. A particularly important problem that causes challenges for these methods is to model systems at finite temperatures ($T > 0$). Some methods have been proposed that extend the methods to that regime and have been shown to work well [56]. However, the size limitations of the calculations limit their applicability to many problems. In later chapters we will introduce methods that will enable us to model very large systems of atoms at finite temperature. We will, however, have to give up the direct calculation of the electronic states and replace them with approximate descriptions of the bonding in materials.

Despite the limitations, density functional calculations have become ubiquitous in materials research, with applications that range from biological systems to alloy development in metals. Here we list just a few examples, showing the range of applicability of these methods.¹⁸

- Structure and thermodynamics: application of *ab initio* methods to calculate structural properties has become routine, as witnessed by a study in which they used DFT to create tables of the ground state energy and structures of 80 binary alloys [75]. A common use of these methods is to calculate thermodynamic properties of systems, one example being a study of intermetallic compounds and solution phases in Sn/Zn [325], and another being a study of the equation of state and elastic properties of a series of metals [31].
- Extreme conditions: another common use of these methods is to calculate properties of materials in regimes for which experimental data might be limited, for example a study of the high-pressure properties of ammonia borane (NH_3BH_3) [202].

¹⁸ The reader is urged to look in the literature for application of electronic structure methods to systems that interest them.

- Defect structure and properties: defect structures are also commonly studied with these methods. One example from a myriad of studies is an examination of the fracture properties of Al/TiN interface [369]. DFT calculations have also been used to shed light on dislocation core structures, for example in aluminum [360].
- Biomaterial applications: biological applications are also now common, an example being a study of how oxygen enters a protein structure [69].
- New materials: electronic structure calculations have also been extremely influential in defining the basic physics of new types of materials. One example is the work of Nicola Spaldin, which ignited interest in the multiferroics [145].

These are just a few examples that indicate the range of applications of electronic structure calculations.

4.9 SUMMARY

In this chapter we introduce the basics of the density functional theory of the electronic structure of solids. After reviewing some of the historical approaches, we describe in some detail the basis of most electronic structure calculations done today, the Kohn-Sham method. We discuss the use of the local-density approximation (LDA) as well as improvements on it through the use of generalized gradient approximations (GGA).

These methods, while not perfect, provide a direct way to calculate the properties of materials with the fewest possible assumptions. They enable us to see where electrons are and let us avoid some of the pitfalls that arise from the use of approximate descriptions of the bonding between atoms, namely the interatomic potential functions described in the next chapter. However, electronic structure methods are computationally costly and thus can only be used for relatively few atoms at a time. Thus, despite their utility, we often must use the approximate methods discussed in later chapters.

Suggested reading

There are many books and articles available that discuss this topic, with a wide range of readability, for example:

- An excellent and comprehensive text by E. Kaxiras, *Atomic and Electronic Structure of Solids* [167].
- A basic description of how to do electronic structure calculations is given in *A Primer in Density Functional Theory* [102].
- A recent text by Martin, *Electronic Structure: Basic Theory and Practical Methods*, is an excellent guide to both basics and applications [219].
- Some older, but very useful, texts include: March, *Self-Consistent Fields in Atoms* [217]; and Parr and Yang, *Density-Functional Theory of Atoms and Molecules* [251].