

Figure 9. Valence density of states of the  $CO + MgO$  system. In the bottom figure, the superposition of the density of states of the two non-interacting systems is reported; the other two figures refer to the vertical adsorption on  $Mg^{2+}$  in the two possible vertical orientations.

9. C. J. Bradley and A. P. Cracknell in *The Mathematical Theory of Symmetry in Solids*, Clarendon Press, Oxford (1972).
10. V. R. Saunders, C. Freyria Fava, R. Dovesi, L. Salasco and C. Roetti, *Mol. Phys.* **77**, 629-665 (1992).
11. F.E. Harris in *Theoretical Chemistry: Advances and Perspectives*, edited by H. Eyring and D. Henderson, Vol. 1 (Academic Press), p. 147.
12. V.R. Saunders in *Computational Techniques in Quantum Chemistry*, edited by G.H.F. Diercksen, B.T. Sutcliffe and A. Veillard (Reidel) p. 347 (1975).
13. C. Pisani, R. Dovesi and C. Roetti, *Lecture Notes in Chemistry*, vol 48, Springer, Heidelberg (1988).

## Ab-Initio Approaches to the Quantum-Mechanical Treatment of periodic Systems

Cesare Pisani

Department of Inorganic, Physical and Materials Chemistry, University of Torino, via Giuria 5, I-10125 Torino, Italy

May 17, 1996

**Summary.** A general overview is proposed of the different strategies that are currently adopted for the *ab initio* study of the electronic structure of periodic systems.

**Key words:** Electron gas – Hartree-Fock (HF) – Electron Correlation – Relativistic Effects – Density-Functional Theory (DFT) – Kohn-Sham (KS) – Local Density Approximation (LDA) – Gradient Corrections (GC) – Self-Interaction Correction (SIC) – Pseudopotentials (PP) – Plane Waves (PW) – Gaussian Type Orbitals (GTO) – Lattice Sums – Self-Consistent-Field (SCF) – Size-consistency – Car-Parrinello (CP) – Montecarlo – KKR – Recursion Method – N-scaling

### 1. Introduction

The purpose of this Chapter is to provide an overview of the methods and computational tools currently adopted in the *ab initio* quantum-mechanical (QM) study of the chemical and physical properties of crystals, and more specifically: a) to make it easier for non-specialists to find their way in a disconcerting variety of proposals, each characterized by a different language, different algorithms, different topical problems; b) to give references for those who want to go deeper in certain subjects; c) to serve as an introduction to subsequent chapters of this book.

Given the chemical composition and the crystalline structure of a periodic system, the aim of *ab initio* computational methods is to calculate its chemical and physical properties as accurately as possible at a reasonable cost, without the need for empirical *a priori* information (although such information can be useful, see section 3.1). In principle, a QM description should be provided for all involved particles, that is, for all electrons and nuclei. In fact, the Born-Oppenheimer

separation is universally adopted: the potential energy surface is first determined, corresponding to the fundamental level of the electronic subsystem for each given nuclear configuration; nuclear motion is studied *a posteriori*, by considering the electronic energy as an external potential [1]. In this Chapter and in most of the subsequent ones, we shall be concerned only with electronic properties, that is, those that depend on the electronic structure of the crystal. This, in turn, is a function of the geometry and the electric charge of the nuclei, which will be taken as fixed and arranged according to a crystalline lattice. The problem of nuclear motion is one of great relevance, because thermodynamic, acoustic and in part, transport properties of the system depend precisely on nuclear vibration or migration. Chapter 12 will discuss some aspects of this problem (see also section 5 of this Chapter).

Preliminary to any QM study of the electronic structure is the choice of the electronic hamiltonian, that is, one must decide which energetic contributions should be included and which expression should be adopted for them. We shall make reference in the following to the non-relativistic electrostatic Hamiltonian  $\hat{H}^{nr}$ . It includes the kinetic energy of electrons, their Coulomb interaction with nuclei and between themselves, as well as the electrostatic interaction between nuclei as a constant term:

$$\hat{H}^{nr} = - \sum_i (\nabla_i^2/2) - \sum_{i,A} Z_A/r_{iA} + \sum_{i>j} 1/r_{ij} + \sum_{A>B} Z_A Z_B / r_{AB} \quad (1)$$

Here and in the following, use is made of atomic units, a.u., see table 1.

**Table 1.** Atomic units.  $e$  and  $m_0$  are the electron charge and rest mass, respectively;  $h$  is the Planck constant.

Quantity	Atomic unit (name)	SI Equivalent
Mass	$m_0$	$9.109610^{-31} \text{ kg}$
Charge	$e$	$1.602210^{-18} \text{ C}$
Angular momentum	$\hbar$	$1.054610^{-34} \text{ Js}$
Length	$a_0 = \hbar^2/(me^2)(\text{Bohr})$	$5.291810^{-11} \text{ m}$
Energy	$E_h = e^2/a_0$ (Hartree)	$4.359810^{-18} \text{ J}$
Time	$t_0 = \hbar/E_h$	$2.418910^{-17} \text{ s}$
Velocity	$a_0/t_0 = E_h/(me^2) \approx c/137$	$2.157910^6 \text{ ms}^{-1}$

Neglecting relativistic effects may be a poor approximation in some instances. The mean square velocity  $v$  of an electron in a  $1s$  orbital is  $Z$  a.u.: for heavy atoms it may be an appreciable fraction of the velocity of light ( $c = 137.036 \text{ a.u.}$ ). In mercury ( $Z = 80$ ),  $v/c \approx 0.58$ , and the corresponding relativistic mass is  $m = 1.23 m_0$ . Since the electron mass appears in the denominator of the Bohr radius, the  $1s$  orbital is contracted by about the same amount. Such important relativistic effects concerning core electrons, indirectly affect valence electrons. In heavy atoms, valence  $s$  and  $p$  orbitals are contracted and stabilized with respect to the non-relativistic case, due to the orthogonality constraints with the orbitals of the innermost shells; on the other hand, valence orbitals of  $d$  and  $f$  type are significantly expanded and destabilized, because of the more effective screening

of the nuclear charge by the  $s$  and  $p$  electrons. These effects can be partially taken into account by using a relativistic treatment for core electrons (see Chapter 9), or by adopting relativistic pseudo-potentials in valence-only calculations (see section 2.4). In addition, the relativistic Dirac hamiltonian contains the spin variable explicitly (for instance, in the so-called spin-orbit coupling term) [2] while in non-relativistic quantum mechanics, spin effects are important only because of the exclusion principle. A recent study of mercury (atom, clusters and bulk) provides an instructive example of the importance of relativistic effects in solid state theory [3]. Neglected relativistic terms may be taken into account *a posteriori* by a perturbative treatment or some other corrective scheme.

The knowledge of the Hamiltonian, and of the total number of electrons contains *in nuce* all information about the electronic structure of the system, both concerning static and time-dependent features. In the following, we shall take the system as neutral, that is, the number of electrons per crystalline unit cell,  $n$ , will equal the sum of the nuclear charges of the translationally inequivalent nuclei. To make progress, in particular to be able to exploit the essential simplification introduced by the periodic nature of the external field, we must specify which aspects of the electronic structure we are mainly interested in. We shall see that the ground state electronic structure (GSES) plays a fundamental rôle in solid state physics; its determination, and the associated properties will be the main object of our interest. A straightforward, though unfeasible means for determining the GSES would be to solve the corresponding Schrödinger equation:

$$\hat{H}^{nr} \Psi_0 = E_0(R) \Psi_0 \quad (2)$$

where the ground state energy,  $E_0(R)$ , corresponds to the lowest eigenvalue for a given nuclear configuration,  $R$ . In fact, it is not only impossible to determine, but also to manipulate such a cumbersome object as  $\Psi_0$  (an antisymmetric function of  $3nN$  spatial coordinates and  $nN$  spin co-ordinates, where  $N$  is the number of crystalline cells, see Chapter 2). It is better to consider simpler functions, which contain as much useful information as possible. One possible choice is to look for simplified forms of  $\Psi_0$ : in particular, in the Hartree-Fock (HF) approach, we approximate  $\Psi_0$  by a Slater determinant, corresponding to the antisymmetrized product of  $nN$  one-electron functions. Other quantities can be chosen to describe the GSES. The one-electron density matrix (or one-matrix):

$$\gamma_0(\mathbf{r}, \mathbf{r}') = nN \int \cdots \int d\mathbf{r}_2 \cdots d\mathbf{r}_{nN} d\sigma_1 \cdots d\sigma_{nN} \times \\ \times [\Psi_0(\mathbf{r}, \sigma_1; \mathbf{r}_2, \sigma_2; \cdots; \mathbf{r}_{nN}, \sigma_{nN}) \Psi_0^*(\mathbf{r}', \sigma_1; \mathbf{r}_2, \sigma_2; \cdots; \mathbf{r}_{nN}, \sigma_{nN})] \quad (3)$$

is a very simple function which contains a lot of important information (see Chapter 14) and is the primary objective of some procedures (see sections 5 and 6). In density functional theory (DFT), the objective is still simpler, namely, the determination of the electron density:  $\rho_0(\mathbf{r}) \equiv \gamma_0(\mathbf{r}, \mathbf{r})$ ; from the knowledge of this function of three spatial coordinates, all ground state properties can be derived, in principle, by means of suitable integrals (see sections 2.2 and 6). The dream of solid state physicists is to determine the self-energy function,  $\Sigma(\mathbf{r}, \mathbf{r}', \epsilon)$ , which would permit all ground state one-particle properties to be obtained, including one-electron excitation energies [4];  $\Sigma(\mathbf{r}, \mathbf{r}', \epsilon)$  can be estimated, for instance, through Hedin's GW-approximation [5].

The main obstacle to the solution of equation 2, or to the determination of some intermediate quantity which contains the desired information, lies in the inter-electronic interaction term,  $\sum_{i>j} 1/r_{ij}$ , which prevents the problem from being factorized into  $nN$  problems for the individual electrons. An essential simplification, therefore, consists in introducing an effective potential for each electron in the field of the others, and substituting a sum of one-electron Hamiltonians ( $\hat{H}^{nr} \approx \sum_i \hat{h}_i$ ) for the full Hamiltonian, at the cost of neglecting some higher order effects to be specified below. This allows us to achieve two advantages simultaneously: firstly, to de-couple the electronic motions; secondly, to introduce, quite naturally, the translational symmetry of the problem. In fact, since we are interested in the GSES, any effective potential must reflect the symmetry of the external potential and therefore be translationally periodic: the methods, language and techniques of band theory can be used (Chapter 2). The following scheme summarizes the concepts introduced up to now (in the lines below, we list the approximations introduced and the effects that are disregarded):

$$\begin{array}{ccccccc}
 \Rightarrow & \text{(fixed nuclei)} & \Rightarrow & \hat{H}_{nr} & \Rightarrow & \{\hat{h}_i\} & \Rightarrow & \text{GSES} \\
 \Downarrow & & \Downarrow & & \Downarrow & & \Downarrow & \\
 \text{nuclear} & & \text{relativistic} & & \text{higher order} & & \text{excited} & \\
 \text{motion} & & \text{effects} & & \text{effects} & & \text{states} & 
 \end{array}
 \quad (4)$$

In the rest of this Chapter, we will concentrate our attention on the last two parts of this scheme. Section 2 is concerned with the formulation of one-electron approximations for the Hamiltonian, starting from the consideration of the prototypical periodic system, the homogeneous electron gas. Two *ab initio* schemes are currently adopted. The first is based on DFT, and its fundamental objective is to derive the correct ground state electron density,  $\rho_0(\mathbf{r})$ , and the exact ground state energy,  $E_0$ ; the second consists of adopting the HF approximation, and aims at obtaining the best single-determinant representation of the ground state wave-function in a variational sense. The fundamental limitation of DF theory is that one must use approximate expressions for the effective potential, since an exact formulation is not available. On the other hand, with the HF approach one is always faced with an electronic correlation problem [1, 6], as is discussed below. The use of pseudo-potentials (PP) to dispense with the treatment of core electrons in one-electron hamiltonians is well established in solid state theory and is summarily treated here.

The intermediate object of our study (the GSES, described either by  $\gamma_0(\mathbf{r}, \mathbf{r}')$ , or by a simplified expression for  $\Psi_0$ ) is some function of the electron coordinates which is usually expressed as a linear combination of a finite set of one-electron functions,  $\phi_\mu(\mathbf{r})$  ( $\mu = 1, \dots, p$ ) or of the products thereof. The choice of the basis set is one of the critical aspects of most computational schemes: the algorithms, the efficiency, the very structure of the program are often intimately related with this choice. A general introduction to the basis set problem in *ab initio* crystalline studies is provided in section 3. More detailed analyses, with reference to specific choices, are contained in Chapters 9, 10 and 11.

Section 4 outlines the standard methods for the determination of the GSES, which are based on the solution of one-electron equations and on the use of band theory. The key features and the critical computational issues are similar for DF and HF schemes and are exposed briefly in parallel.

Non-standard methods which have relevance to solid state studies and are not considered in this book, are summarily treated in section 5 for information. The general problem of the scaling of the computational cost with the size of the irreducible part of the crystal is also considered.

Section 6 serves as an introduction to the last part of this book, concerning the calculation of crystalline properties.

## 2. One-Electron Hamiltonians

### 2.1. Introduction: The electron gas problem

The problem of the homogeneous electron gas (a system of interacting electrons in a uniform sea of compensating positive charge) has traditionally played a crucial rôle in solid state theory, but its importance has grown dramatically with the advent and the extraordinary success of DF theory. The discussion of the merits and limitations of one-electron Hamiltonians in solid state physics can begin with some of the essential results concerning this ideal system. A concise account of the subject is given in chapter 2 of Fulde's book [6]; a clear and detailed treatment (except for the more recent studies) can be found in Raimès' book [7].

The only parameter which characterizes the electron gas is its uniform density,  $\rho$  (number of electrons per unit volume), or equivalently, the radius of the Wigner-Seitz sphere,  $r_s = [(4/3)\pi\rho]^{-1/3}$ . One often speaks of *low density* or *high density* electron gas according to whether  $r_s > 1$  or  $r_s < 1$  (for instance, for valence electrons in sodium  $r_s = 3.9$ , for core electrons in carbon  $r_s = 0.22$ ; as always, a.u. are used). Two one-electron approximations are applied to solve the non-relativistic problem of equation 2. The *free-electron* approach takes into account the fact that classically, the electrostatic self-interactions of the electron gas and of the positive background and their mutual interaction cancel exactly. The only term which is left is the sum of the kinetic energy terms for the individual electrons. In addition, an exchange term appears in the HF approximation, which takes into account the antisymmetry of the wave-function (see section 2.3). The problem then becomes an elementary one for the individual electrons:

$$\left[-\frac{1}{2}\nabla^2 + (v_x)_i\right]\psi_i(\mathbf{r}) = \varepsilon_i\psi_i(\mathbf{r}) \quad (5)$$

In this equation, the exchange term,  $(v_x)_i$ , is zero for the free-electron gas approach; in the HF scheme, it is a constant which may depend on  $i$  but is independent of  $\mathbf{r}$  for symmetry reasons. The eigenvectors are then plane waves:  $\psi_i(\mathbf{r}) \equiv \psi(\mathbf{r}; \mathbf{k}) = \Omega^{-1/2} \exp(i\mathbf{k} \cdot \mathbf{r})$ , with eigenvalues  $\varepsilon_i \equiv \varepsilon(\mathbf{k}) = \frac{1}{2}k^2 + v_x(\mathbf{k})$  (for symmetry reasons,  $\varepsilon_i$  cannot depend on the direction of  $\mathbf{k}$ ). According to the Aufbau principle, the occupied orbitals fill a sphere in  $\mathbf{k}$ -space, centered in  $\mathbf{k} = 0$  of radius  $k_F = 1/(\alpha r_s)$ , with  $\alpha = [4/(9\pi)]^{1/3} = 0.521$ . Since the expression of the occupied orbitals is known, the HF value of  $v_x(\mathbf{k})$  can be calculated and it turns out to be:

$$v_x(k) = \frac{k_F}{2\pi} \left[ 2 + \frac{k_F^2 - k^2}{k \cdot k_F} \cdot \log \left| \frac{k + k_F}{k - k_F} \right| \right] \equiv \frac{k_F}{2\pi} F\left(\frac{k}{k_F}\right) \quad (6)$$

The GS energy per electron is easily calculated in the two cases. For the free-electron approximation, it is purely kinetic:

$$\varepsilon_{kin}(\rho) = (3/10)/(\alpha^2 r_s^2) = (3/10)[3\pi^2]^{2/3} \quad (7)$$

In the HF case, it contains, in addition, an exchange term:  $\varepsilon_{HF}(\rho) = \varepsilon_{kin}(\rho) + \varepsilon_x(\rho)$ , with:

$$\varepsilon_x(\rho) = -3/(4\pi\alpha r_s) = -(3/4)/(3\rho/\pi)^{1/3}. \quad (8)$$

The results just presented have had great historical importance in creating a bad reputation for the HF approximation among solid state physicists. Consider first the density of states (DOS), that is the number of one-electron states per unit energy interval and per unit volume. It is tempting to use the information above to calculate the distribution of electronic excitation energies, an excited state being created by letting one or more electrons pass from an occupied level,  $\varepsilon$ , to an empty one,  $\varepsilon'$ , with a corresponding energy cost:  $\Delta\varepsilon = \varepsilon' - \varepsilon$ . This assumption is strictly valid within the free-electron approximation, and finds a loose justification in Koopmans' theorem for the HF model (see section 2.3). Using the free-electron levels  $[\varepsilon(k) = k^2/2]$  gives satisfactory agreement with experiment for many electronic properties (low temperature specific heat, electrical conductivity, spin susceptibility among others) of alkali metals such as lithium or sodium, for which the electron gas model is best justified. The addition of the exchange correction,  $v_x(k)$  (equation 6), instead of improving the agreement, introduces serious discrepancies with respect to experiment. First of all, the energy spread of occupied levels is too large. More fundamentally, it is easily seen that the derivative with respect to energy of the  $F(k/k_F)$  factor in equation 6 exhibits a logarithmic divergence when  $k = k_F$ : this corresponds to zero density of states at the Fermi level, obviously contrary to experimental evidence. This paradoxical result is not inherent in the homogeneous electron gas model, but has been shown to be a necessary outcome of the HF equations when applied to metallic systems [8, 9]. It is the result of neglecting correlation effects in the HF approximation, which are essential for screening the  $r^{-1}$  dependence of the Coulomb interactions in a metal.

Correlation effects are of decisive importance also for the cohesive energy of the electron gas. In the HF expression for the specific energy:  $\varepsilon_{HF}(\rho) = \varepsilon_{kin}(\rho) + \varepsilon_x(\rho)$ , the exchange term makes the electron gas stable for  $r_s > 2.41$ , but in all cases, the calculated cohesion energy is low with respect to the experimental cohesion energy of alkali metals. A great deal of theoretical and computational work has been devoted to obtaining a good estimate of the correlation energy for the electron gas. The many-body study by Gell-Mann and Bruckner for high densities [10], and the more recent quantum-Montecarlo results of Ceperley and Alder for low densities (see section 5) [11] are particularly important. The calculated data have been recently parameterized by Perdew and Zunger [12], in order to obtain an analytic expression, which is useful for DF calculations (see section 2.2):

$$\begin{aligned} \varepsilon_c(\rho) &= 0.0311 \log r_s - 0.048 + 0.0020 r_s \log r_s - 0.0116 r_s \quad [\text{for } r_s < 1] \\ \varepsilon_c(\rho) &= -0.1423/(1 + 1.0529\sqrt{r_s} + 0.3334r_s) \quad [\text{for } r_s > 1] \end{aligned} \quad (9)$$

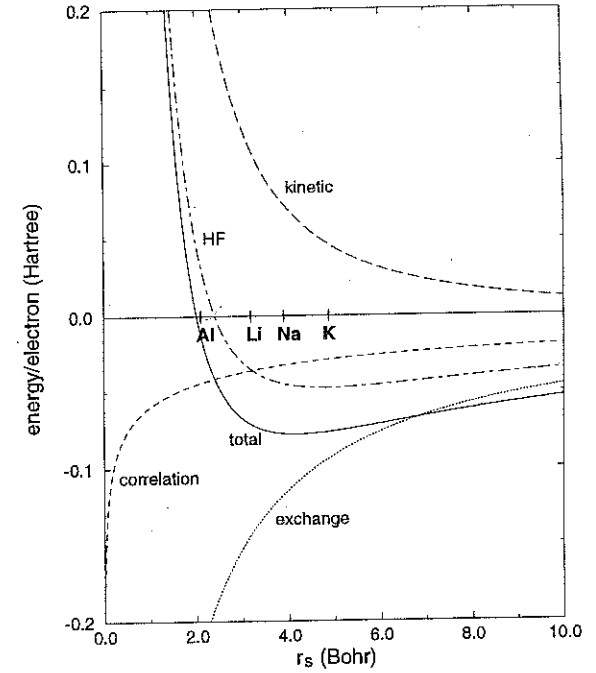


Figure 1. Contributions to the energy per electron for the homogeneous electron gas, as a function of the Wigner-Seitz radius,  $r_s$ . The  $r_s$  values of valence electrons of a few simple metals are reported.

Figure 1 shows the importance of correlation corrections in the region of electron densities of chemical importance. The effect of external fields on the properties of the electron gas is a subject of great interest, still intensively studied, especially for its applications in DF work (see section 2.2).

## 2.2. Density functional theory

DF techniques are based on two theorems, proved by Hohenberg and Kohn in 1964 [13], and on a computational scheme proposed by Kohn and Sham (KS) the following year [14].

The first theorem states that a given ground state (GS) electron density,  $\rho_0(\mathbf{r})$ , cannot arise from two different external potentials, unless the two differ by a constant. Hence, the GSES ( $\Psi_0$  and all derived GS observables, in particular  $E_0$ ) is uniquely determined by  $\rho_0(\mathbf{r})$ .

The second theorem establishes a variational criterion for determining  $\rho_0(\mathbf{r})$  and  $E_0$  together for a given external potential created, for example, by a set

of nuclear charges:  $V(\mathbf{r}) = -\sum_A Z_A/|\mathbf{r} - \mathbf{r}_A|$ .  $E_0$  is found by minimizing an expression of the form:

$$\begin{aligned} E[\rho(\mathbf{r})] &= \int \rho(\mathbf{r})V(\mathbf{r})d\mathbf{r} + \frac{1}{2} \int [\rho(\mathbf{r})\rho(\mathbf{r}')]/|\mathbf{r} - \mathbf{r}'| d\mathbf{r} d\mathbf{r}' + \\ &+ \int \rho(\mathbf{r})g(\mathbf{r};[\rho]) d\mathbf{r} + \sum_{A>B} Z_A Z_B / r_{AB} \\ &= E_{Ext} + E_J + E_{kxc} + E_{RepNuc} \end{aligned} \quad (10)$$

with respect to an arbitrary function,  $\rho(\mathbf{r})$ , which represents an  $n$ -electron density. The minimum is obtained precisely in correspondence to  $\rho_0(\mathbf{r})$ . In this expression, the first term gives the interaction energy with the external potential, the second term is the classical self-interaction of the charge density  $\rho(\mathbf{r})$ , while the last term is the usual internuclear repulsion. The third term,  $E_{kxc}$ , contains the  $g(\mathbf{r})$  function, which may be interpreted as a density of kinetic, exchange and correlation energy per unit electron density at  $\mathbf{r}$ .  $g(\mathbf{r})$  is known to be a universal *functional* of the density  $\rho$  (that is, it uniquely depends on the *whole* charge distribution). Its analytic form is not known, but different approximations are used. A local density approximation (LDA) consists of making  $g(\mathbf{r})$  depend only on the value of  $\rho$  at  $\mathbf{r}$ . Among LDAs, the historically important Thomas-Fermi approximation [15] is obtained by giving  $g(\mathbf{r})$  the value of the specific kinetic energy of an electron gas with uniform density,  $\rho(\mathbf{r})$ :  $g(\mathbf{r}) \approx (3/10)[3\pi^2\rho(\mathbf{r})]^{2/3}$ .

The most critical aspect of  $g(\mathbf{r};[\rho])$ , however, is precisely that it contains the kinetic energy term: the Thomas-Fermi approximation is not very accurate but it is difficult to improve on it. The KS method is based on the two theorems above but permits us to calculate explicitly the most important part of the kinetic energy. In fact, it consists of solving a parallel problem for a system of  $n$  pseudo-independent electrons, characterized by the same GS density as the actual system. For the pseudo-system, the pseudo-kinetic energy  $E_{kps}$  can be explicitly calculated and the problem is solved according to the following self-consistent field (SCF) procedure:

- a) Find the  $n/2$  lowest eigenvalues and the corresponding eigenfunctions (*orbitals*) of a one-electron effective Hamiltonian:

$$\hat{h}_{eff} \psi_i(\mathbf{r}) \equiv [-\frac{1}{2}\nabla^2 + V_{eff}(\mathbf{r})]\psi_i(\mathbf{r}) = \varepsilon_i \psi_i(\mathbf{r}); \quad (11)$$

- b) Calculate the density:

$$\rho(\mathbf{r}) = 2 \sum_i |\psi_i(\mathbf{r})|^2; \quad (12)$$

- c) Re-calculate  $V_{eff}(\mathbf{r})$  as a functional of  $\rho(\mathbf{r})$ :

$$V_{eff}(\mathbf{r}) = V(\mathbf{r}) + \int \rho(\mathbf{r}')/|\mathbf{r} - \mathbf{r}'| d\mathbf{r}' + \mu_{xc}(\mathbf{r};[\rho]); \quad (13)$$

if self-consistency is not reached, go to step a).

- d) At self-consistency,  $\rho(\mathbf{r}) = \rho_0(\mathbf{r})$  (GS density) and the GS energy can be calculated:

$$\begin{aligned} E_0 &= E_{kps} + \int \rho_0(\mathbf{r})V(\mathbf{r})d\mathbf{r} + \frac{1}{2} \int [\rho_0(\mathbf{r})\rho_0(\mathbf{r}')]/|\mathbf{r} - \mathbf{r}'| d\mathbf{r} d\mathbf{r}' + \\ &+ E_{xc} + E_{RepNuc} \\ E_{kps} &= \sum_i \int \psi_i(\mathbf{r})^* (-\frac{1}{2}\nabla^2) \psi_i(\mathbf{r}) d\mathbf{r} \\ E_{xc} &= \int \rho_0(\mathbf{r})\varepsilon_{xc}(\mathbf{r};[\rho_0])d\mathbf{r} \end{aligned} \quad (14)$$

The two universal functionals  $\mu_{xc}(\mathbf{r};[\rho])$  in equation 13 (effective exchange-correlation potential) and  $\varepsilon_{xc}(\mathbf{r};[\rho])$  in equation 14 (specific exchange-correlation energy) are related to each other by a *functional derivative* equation:

$$\mu_{xc}(\mathbf{r};[\rho]) = \delta E_{xc}([\rho])/\delta\rho(\mathbf{r}) \equiv \delta\left\{ \int \rho(\mathbf{r}')\varepsilon_{xc}(\mathbf{r}';[\rho])d\mathbf{r}' \right\}/\delta\rho(\mathbf{r}) \quad (15)$$

The following can be noted:

- The beauty and the originality of the KS scheme lies in the fact that its aim is not to give an approximation to  $\Psi_0$ , but to yield the exact GS energy  $E_0$  and the exact GS electron density,  $\rho_0(\mathbf{r})$ , directly. In this sense, it is important to observe that the pseudo-wave-function, defined as an antisymmetrized product of the pseudo-spin orbitals:  $\Psi_{KS} = \hat{A} [\prod_i (\psi_i(\mathbf{r}_{2i-1})\alpha(\sigma_{2i-1})\psi_i(\mathbf{r}_{2i})\beta(\sigma_{2i}))]$ , has nothing to do, in principle at least, with the true  $\Psi_0$ , except that it defines the same GS density. This is equally true for the pseudo-eigenvalues  $\varepsilon_i$ , which should not be used as indicative of one-electron excitation energies. On the other hand, it has been customary in solid state applications to adopt such an inclusive interpretation of the results of the KS equations: this attitude has some theoretical ground (see section 6) and is justified by many remarkable successes. Part of the theoretical work in the field of DF theory aims at defining an effective potential which generates pseudo-eigenvalues  $\varepsilon_i$  more strictly related to one-particle excitations [16].
- The crucial problem in the KS method is the definition of the exchange-correlation functional (it should be noted that this definition is not fully appropriate, because  $E_{xc}$  must contain, in addition to the exchange and correlation energies, a correction to the pseudo-kinetic energy  $E_{kps}$ ). For this purpose, reference is most often made to the results for the homogeneous electron gas. The LDA is by far the most frequent choice:

$$\varepsilon_{xc}(\mathbf{r};[\rho]) \approx \varepsilon_x(\rho(\mathbf{r})) + \varepsilon_c(\rho(\mathbf{r}))$$

where  $\varepsilon_x(\rho)$  and  $\varepsilon_c(\rho)$  are given by expressions (8) and (9) (when applied to the electron gas, LDA-DF theory is exact but tautological). In the LDA, the relationship between  $\mu_{xc}$  and  $\varepsilon_{xc}$  simplifies to:  $\mu_{xc}(\rho) = d[\rho \varepsilon_{xc}(\rho)]/d\rho$ . The celebrated X- $\alpha$  method of Slater [17] can be interpreted as an anticipation of the KS-LDA equations, with only the exchange LDA potential included, multiplied by an empirical parameter ( $3\alpha/2$ ).

- Given a fairly regular density, any functional  $\alpha(\mathbf{r}; [\rho])$  is a *function* of  $\rho$  and all its derivatives at  $\mathbf{r}$ . The LDA can be viewed as the zero term in a gradient expansion of the exact functionals [13]. In recent years, the use of gradient-corrected (GC) functionals has become more and more wide-spread, which corresponds to going a step farther in the gradient series. In these cases, the value of the  $\varepsilon_{xc}$  functional depends both on the value of the density at  $\mathbf{r}$ ,  $\rho(\mathbf{r})$ , as well as on the magnitude of the density gradient there,  $|\nabla\rho(\mathbf{r})|$ . For the parameterization of such functionals, reference is made to results for the inhomogeneous electron gas [18]. On the other hand, the convergence properties of the gradient series are unknown and making real progress by adding more of the neglected, higher-order terms (see equation 4) seems a hopeless task.
- Up to now, we have considered only closed-shell systems. The extension of DF theory to spin polarized systems is quite natural [19, 12]. One must define an electron density for each spin state:  $\rho^\uparrow$ ,  $\rho^\downarrow$  or equivalently, a total density,  $\rho = \rho^\uparrow + \rho^\downarrow$  and a spin electron density  $m = \rho^\uparrow - \rho^\downarrow$ . All functionals, for instance  $\varepsilon_{xc}$ , will depend on both densities:  $\varepsilon_{xc} \equiv \varepsilon_{xc}(\mathbf{r}; [\rho, m])$ . The discussion of this class of important problems goes beyond the scope of this chapter.
- The electron-electron Coulomb interaction energy (see equation 10) contains a spurious contribution arising from the interaction of each electron with itself. This should be exactly cancelled by an exchange term of opposite sign but this does not happen in standard applications of DF theory, due to the approximations involved in the calculation of  $E_{xc}$ . There is a need for a self-interaction-correction (SIC) to the KS equations, especially in their application to molecular systems. Different SIC methods have been proposed, which consist essentially of subtracting, from the total energy expression, the sum of the self-interactions in the individual pseudo-orbitals [12, 20]. For a recent critical account, see references [21].
- The history of DF applications in solid state physics is too rich to be even summarily sketched here; some examples will be given in subsequent Chapters of this book. The pioneering work of M.L. Cohen and collaborators at Berkeley should however be mentioned, who produced in the seventies and eighties an ample and convincing documentation of the power, generality and reliability of the *ab initio* DF treatment of crystalline systems [22].

### 2.3. Hartree-Fock theory and the electron correlation problem

The HF equations are homologous to the KS ones, in the sense that we are again looking for a set of  $n/2$  orbitals as eigenfunctions of a one-electron Hamiltonian, the Fock Hamiltonian  $\hat{f}$  (here, reference is made to a closed-shell system; the extension to spin-polarized systems is considered in Chapter 6):

$$\begin{aligned}\hat{f}\psi_i(\mathbf{r}) &\equiv [-(1/2)\nabla^2 + v(\mathbf{r}) + \int d\mathbf{r}' \rho(\mathbf{r}')/|\mathbf{r} - \mathbf{r}'|] \psi_i(\mathbf{r}) + \\ &\quad - \sum_j \int d\mathbf{r}' [\psi_j(\mathbf{r}')^* \psi_i(\mathbf{r}')] / |\mathbf{r} - \mathbf{r}'| \psi_j(\mathbf{r}) = \\ &= \varepsilon_i \psi_i(\mathbf{r})\end{aligned}\quad (16)$$

The first three terms in the Fock Hamiltonian coincide with the corresponding KS ones (the expression for the electron density is the same). However, instead of the effective exchange-correlation potential,  $\mu_{xc}(\mathbf{r}; [\rho])$ , here we have the “exact” non-local exchange operator (the fourth term in equation 16). Again, the HF equations must be solved through an SCF procedure, corresponding to steps a-d of section 2.2, since both the Coulomb and the exchange operator depend on the set of functions,  $\psi_i(\mathbf{r})$ . The essential difference, with respect to the KS approach, is that the antisymmetrized product of the Fock spin orbitals:  $\Psi_{HF} = \hat{A}[\prod_i (\psi_i(\mathbf{r}_{2i-1})\alpha(\sigma_{2i-1})\psi_i(\mathbf{r}_{2i})\beta(\sigma_{2i}))]$  is an approximation of the GS wave function: more precisely, it is the best single-determinant approximation (in a variational sense) of  $\Psi_0$ . Similar comments apply concerning the one-matrix (see Chapter 14). The HF energy,  $E_{HF}$ , is defined as the expectation value of the non-relativistic Hamiltonian for  $\Psi_{HF}$ :

$$\begin{aligned}E_{HF} &= \langle \Psi_{HF} | H^{nr} | \Psi_{HF} \rangle \\ &= 2 \sum_i \int d\mathbf{r} \{ \psi_i(\mathbf{r})^* [-\nabla^2/2 + V(\mathbf{r})] \psi_i(\mathbf{r}) \} + \\ &\quad + \sum_{ij} \int d\mathbf{r} d\mathbf{r}' [\psi_i(\mathbf{r})^* \psi_i(\mathbf{r})] [\psi_j(\mathbf{r}')^* \psi_j(\mathbf{r}')^*] / |\mathbf{r} - \mathbf{r}'| + \\ &\quad - 1/2 \sum_{ij} \int d\mathbf{r} d\mathbf{r}' [\psi_i(\mathbf{r})^* \psi_j(\mathbf{r})] [\psi_j(\mathbf{r}')^* \psi_i(\mathbf{r}')^*] / |\mathbf{r} - \mathbf{r}'| \quad (17)\end{aligned}$$

Here are a few comments on the HF method:

- The non-local exchange term makes the HF equations more difficult to solve than the KS ones, where  $\mu_{xc}(\mathbf{r}; [\rho])$  is simply a multiplicative operator, no matter how complicated its determination. This is specially true for crystalline systems, as we shall see in section 4. If one adds the fact that HF performs poorly for the electron gas, the simplest of all periodic systems (section 2.1), the lack of popularity of this approximation among solid state physicists would appear well justified. This contrasts with its almost universal acceptance in molecular studies. Early attempts to formulate and implement HF computational schemes for periodic systems have had scant success, in spite of the high quality of those studies [23, 24, 25, 26]. Only in recent years has it become possible to formulate a fair judgment about the usefulness of the HF approach in solid state physics, since the advent of powerful computers and of general purpose computational schemes [27, 28], which have allowed us to assess its performance for a variety of systems (see section 7).
- At variance with DF theory, one-electron HF eigenvalues *can* be attributed a physical meaning. If one creates an  $(n-1)$ -electron state, represented by a single-determinant  $\Psi'$  which may be obtained from  $\Psi_{HF}$  by eliminating a spin-orbital with eigenvalue  $\varepsilon_i$ , it is easily shown that  $\varepsilon_i = \Delta E \equiv \langle \Psi_{HF} | H^{nr} | \Psi_{HF} \rangle - \langle \Psi' | H^{nr} | \Psi' \rangle$  (Koopmans' theorem) [1].  $\Delta E$  (or  $\varepsilon_i$ ) can be considered as an ionization potential of the system, if one assumes that the removal of the electron does not affect the other occupied orbitals (the frozen-orbital approximation). The neglect of relaxation tends to produce ionization potentials that are too positive; correlation effects, whose importance grows with the number of electrons, correct partly for this error. For molecules,

Koopmans' ionization potentials are usually in reasonable agreement with experiment [1].

- As a by-product of the solution of the HF equations, virtual orbitals and energy levels are obtained, corresponding to solutions of equation 11 with eigenvalues higher than those of occupied orbitals. Using Koopmans' theorem, one can estimate electron affinities, by adding an electron in a virtual orbital. In this case, however, the correlation correction adds to the relaxation error, and electron affinities are far too negative. Similarly, it is incorrect to use one-electron HF eigenvalues to calculate the excitation energy resulting from the promotion of an electron from an occupied ( $\psi_i$ ) to a virtual ( $\psi'_j$ ) energy level as  $\Delta E = \varepsilon'_j - \varepsilon_i$ . Therefore, HF band structures are not suitable for the estimation of optical excitations in crystals: the HF main gap for insulators and semi-conductors is usually about twice as large as that given by experiment [27]. On the other hand, it is relatively easy to correct HF band-structures to reproduce the spectrum of optical excitations, by using, for instance, the GW approximation [5, 29, 30].
- The main limitation of the HF solution lies in the neglect of correlation effects. Consideration of the interaction of each electron with the average field of the others leads to an over-estimation of electron-electron repulsions: in fact, electrons move so as to keep apart from each other (otherwise stated, each electron carries along a *correlation hole*). The recent book by Fulde discusses the correlation problem in crystalline studies and can be used as a fundamental reference [6].
- If we are interested only in an estimate of the correlation energy,  $E_c = E_0 - E_{HF}$ , simple and powerful techniques are available, based on the precepts of DF theory. As was shown in the preceding sub-section, it is possible to define an LDA or a GC expression for *correlation-only* functionals. In Chapter 11, indications are given of the efficiency of these techniques.
- While in molecular quantum chemistry the electron correlation problem has been studied in detail for several decades now and many different powerful techniques have been devised to cope with it, the state of the art in solid state theory is far less advanced, not only because of the difficulty of obtaining good crystalline HF wave-functions. Most standard molecular schemes use the set of occupied and virtual orbitals as basic ingredients. In crystals, these are de-localized and correspond to a continuum of levels, which makes their treatment awkward. Furthermore, a correlation scheme must be *size-consistent* in order to be transferable to solid state applications. The size-consistency requirement can be stated as follows. The application of the scheme to a system composed of  $N$  identical and independent subsystems, must give exactly  $N$  times the correlation energy obtained for the individual system. The most standard of all correlation schemes, the variational Configuration-Interaction technique, restrained to single and double excitations (SD-CI) does not satisfy this requirement. The reason is easily understood: the product of the wave-functions for the individual sub-systems, each containing SD excitations, would indeed give  $N$  times the individual energy, but contains excitations involving up to  $2N$  electrons. Among size-consistent schemes which are beginning to be transferred from molecular to solid state applications, we note the coupled-cluster (CC) technique [31], the Moeller-Plesset perturbative series and the Green-function many-body approach, which is particularly

appropriate for correcting the HF band-structure [30]. The *local* approach, which is a variant of the CC method specially designed for crystals, is amply illustrated in Fulde's book [6]. Finally, if one does not impose symmetry restrictions on the spin component of the wave-function and looks for unrestricted HF solutions, one often obtains a lower energy by separating spin-up from spin-down electrons in space; interesting results have been obtained for simple metals [32].

- Fulde's comment about the value of HF computations for solids can be reported in conclusion [6]:

"Provided that the correlations are not too strong, HF calculations are a good starting point, and allow solids and molecules to be treated the same way and with the same accuracy. The development of quantum chemistry has proven that ab initio calculations based on controlled approximations capable of systematic improvement have made simpler computational schemes based on uncontrolled simplifications obsolete. Whether or not the same will eventually hold true for solid-state theory remains to be seen."

## 2.4. Pseudo-potentials

In many respects, core electrons are unimportant for determining the stability, structure and low-energy response properties of molecules and of crystals. It is a well-established practice to modify the one-electron part of the Hamiltonian by replacing the bare nuclear attraction with a PP operator,  $\hat{V}_{ps}$ , which permits us to restrict the calculation to valence electrons, for both types of system. This choice is necessary in practice when the system includes heavy atoms ( $Z > 20$ ). Consider a system comprising a set of nuclei  $A$  and  $n$  electrons, of which  $n'$  are valence electrons. We can write equations of the form:

$$[-\nabla^2/2 + \int \rho'(\mathbf{r}')/|\mathbf{r} - \mathbf{r}'|d\mathbf{r}' + \mu_{xc}(\mathbf{r}; [\rho']) + \sum_A \hat{V}_{psA}] \psi'_i(\mathbf{r}) = \varepsilon'_i \psi'_i(\mathbf{r}) \quad (18)$$

which coincide with the DF-KS equations 11-13, except for the substitution of  $V(\mathbf{r})$  by the PP operator:  $\hat{V}_{ps} = \sum_A \hat{V}_{psA}$  (a parallel procedure can be followed in the HF case). Primed symbols have been introduced to indicate that the SCF solution of these equations is limited to valence electrons; that is, we will consider only the  $n'/2$  eigenfunctions corresponding to the lowest eigenvalues, and use them to calculate the valence density,  $\rho'(\mathbf{r})$ , which in turn is used to define the electrostatic, exchange and correlation interactions between valence electrons. The PP operator must reproduce screened nuclear attractions, but must also account somehow for the Pauli exclusion principle, which requires that valence orbitals are orthogonal to core ones. Consider an atom  $A$  at  $\mathbf{A}$ , with  $n_{cA}$  core electrons, whose highest angular quantum number is  $L$ . Suppose  $L = 1$ , that is, the core contains only  $s$  and  $p$  electrons. At long range, we must have:  $\hat{V}_{psA} = -(Z_A - n_{cA})/r_A$ , with  $r_A = |\mathbf{r} - \mathbf{A}|$ . At short range,  $\hat{V}_{psA}$  must act differently on functions of  $s$  and  $p$  symmetry, while it operates equivalently on functions of higher angular quantum number, that is:



$$\hat{V}_{ps} = -(Z_A - n_{cA})/r_A + \sum_{\ell=0}^L U_{\ell}^{s-r}(r_A) \left\{ \sum_{m=-\ell}^{\ell} |\ell m\rangle \langle \ell m| \right\}_A + W^{s-r}(r_A) \quad (19)$$

$W^{s-r}(r_A)$  and  $U_{\ell}^{s-r}(r_A)$  are short-range functions; the term in braces is a projector, which makes  $U_{\ell}^{s-r}$  act only on functions which have  $\ell$ -symmetry with respect to  $\mathbf{r}_A$ . In establishing the explicit form of  $\hat{V}_{psA}$ , a number of characteristics are sought (not all of them can be optimally satisfied):

- i) Pseudo-valence eigenvalues,  $\varepsilon'_i$ , should coincide with the true ones,  $\varepsilon_i$ ;
- ii) Pseudo-orbitals,  $\psi'_i(\mathbf{r})$ , should resemble as closely as possible the true ones,  $\psi_i(\mathbf{r})$ , in an external region as well as being smooth and node-less in the core region;
- iii) Pseudo-orbitals,  $\psi'_i(\mathbf{r})$ , should be properly normalized;
- iv) The functional form of the PPs, though preserving non-local character, should be designed so as to simplify as far as possible their use in computations;
- v) Explicit expressions for the functions  $W^{s-r}(r_A)$  and  $U_{\ell}^{s-r}(r_A)$  should be provided for the different atomic species, to be used independently of the environment;
- vi) For heavy atoms, relativistic effects should be taken into account (see Introduction): that is, the form of the PP should be derived from relativistic KS or HF atomic calculations.

The performance of PP techniques in solid state physics is usually very good, except for some critical cases (for instance, if core relaxation effects are important, which may occur when simulating very high pressures, or when the electron configuration in the crystalline environment is very different from that of the isolated atom) [22, 33]. Among PPs designed for KS-LDA calculations, the norm-conserving ones tabulated for all atoms by Bachelet et al. [34] are perhaps the most popular; *ultrasoft* PPs, which ensure the very smooth behaviour of the pseudo-valence orbitals in the core region are useful in applications where plane waves are used as a basis set (see section 3.2 and Chapter 10) [35]. With regard to PPs for HF calculations, we may mention those introduced by the Toulouse group [36] twenty years ago, which are still widely used, and the more recent ones designed by Hay and Wadt [37]. In all these cases, relativistic corrections are added for heavy atoms.

### 3. Analytic representation of Wave-function and Operators

#### 3.1. Introductory remarks

There is a huge amount of information contained in the GSES and associated operators. An efficient computational scheme must be able to manipulate this information economically with regard to CPU time, storage requirements and the number of input/output operations. If possible, the algorithms should be simple and general enough to make accumulation and transfer of know-how from one program to another easier. Linear representations of one-electron functions are usually adopted:

$$f(\mathbf{r}) = \sum_{\mu=1}^p c_{\mu} \phi_{\mu}(\mathbf{r}) \quad (20)$$

With few exceptions, the basic ingredients in solid state applications for the construction of the basis-set functions,  $\phi_{\mu}(\mathbf{r})$ , are plane-waves (PW) and/or Gaussian type orbitals (GTO), and they are the only ones that are considered in the following. The basic advantage obtained from the use of PWs and GTOs is related to the fact that they make the computation of integrals in direct and/or reciprocal space very easy. If *numerical techniques* are used to perform integrals throughout the calculation, much greater freedom is possible in the choice of the representative functions: these may include Slater type orbitals (see below), PWs, numerically tabulated orbitals or any combination of these. A KS computational scheme based on a numerical approach has been prepared recently by Baerends and collaborators [38]. All types of symmetry and periodicity in one, two and three dimensions can be handled. The efficiency of the scheme in handling complicated structures is still to be assessed, however.

When using linear representations, the problem of the errors related to the use of an incomplete expansion set arises. There is the need of making such errors as small as possible, while using manageable basis sets. Each application requires a careful analysis: the art of devising good basis sets is a very important one, and is based on experience and competence (see Chapter 8). In this respect, even *ab initio* methods can profit from *empirical knowledge*. One of the advantages that are obtained from the use of Baerends' numerical approach is that the problem of basis set incompleteness, which often plagues schemes based on PWs and GTOs, becomes almost irrelevant.

#### 3.2. Plane waves

PWs are, in a sense, the ideal basis functions for a periodic system. If one denotes the general wave-vector as  $\mathbf{k}$  and the translationally equivalent vector in the first Brillouin Zone (BZ) as  $\boldsymbol{\kappa}$ , we can write:

$$\phi_{\mathbf{k}}(\mathbf{r}) = \Omega^{-1/2} \exp(i\mathbf{k} \cdot \mathbf{r}) = \Omega^{-1/2} \exp[i(\boldsymbol{\kappa} + \mathbf{K}) \cdot \mathbf{r}] \equiv \phi_{\mathbf{K}}(\mathbf{r}; \boldsymbol{\kappa}) \quad (21)$$

The last symbol indicates that the general PW is a Bloch function (BF) associated with the point  $\boldsymbol{\kappa}$  within the BZ and labelled with a discrete index  $\mathbf{K}$ , corresponding to a vector of the reciprocal lattice (see Chapter 2, equation 41). PWs are an orthonormal, complete set: any function belonging to the class of continuous, normalizable functions (which are those of interest in QM) can be expanded with arbitrary precision in the PW set. The set is universal, in the sense that it depends neither on where atoms are located within the cell nor on their nature. As we shall see in section 5, this characteristic is particularly valuable in Car-Parrinello calculations, where nuclear positions are continually changing.

In practice, one must use a finite number of PWs: how many are necessary? Normally one chooses a *cutoff energy*,  $T$ , that is, for each  $\boldsymbol{\kappa}$  in the BZ, one includes all PWs with  $(\boldsymbol{\kappa} + \mathbf{K})^2 < T = X^2$  ( $X$  is the radius of the sphere centered at the origin in reciprocal space which includes all selected PWs) in the expansion set. In typical calculations,  $T$  is of the order of 20 Ry (for historical



reasons, cutoffs are usually measured in Rydbergs: 1 Ry = 0.5 a.u.). If the volume,  $V$ , of the unit cell in direct space is 1000, say, then the volume of the Brillouin zone is  $V_{BZ} = 8\pi^3/V = 0.248$  and the number of  $\mathbf{K}$  vectors, hence of PWs,  $\phi_{\mathbf{K}}(\mathbf{r}; \kappa)$ , contained in the sphere is  $(4\pi X^3/3)/V_{BZ} = 534$ . While this is not a small number of basis functions, the finest detail that can be revealed in direct space with such a set is of the order of  $1/X = 0.58$ , about half the radius of the hydrogen atom. This may be a reasonable figure if we are interested only in valence electrons, but is totally unacceptable for the description of core electrons. For a factor ten in detail, which would still be insufficient except for very light atoms, the number of PWs per  $\kappa$  point should increase by a factor 1000! Unless one *augments* the PW set with some additional function (see section 5 and Chapter 9) the use of PPs is mandatory. In practice, ultrasoft PPs are preferred for obvious reasons (section 2.3).

The use of PWs makes most algebraic manipulations very simple and it is convenient to perform all calculations in the momentum representation. For example, consider the potential  $v_H(\mathbf{r}) = \int \rho'(\mathbf{r}')/|\mathbf{r} - \mathbf{r}'| d\mathbf{r}'$  generated at the general point  $\mathbf{r}$  by the distribution of pseudo-valence electrons, which is one of the components of the KS Hamiltonian (equation 18). Since  $\rho'(\mathbf{r})$  is a periodic function in direct space, so is  $v_H(\mathbf{r})$  and we can write, after setting to zero the average value of  $v_H(\mathbf{r})$ :

$$\rho'(\mathbf{r}) = \sum_{\mathbf{K}} \rho_{\mathbf{K}} \exp(i\mathbf{K} \cdot \mathbf{r}) \quad \text{and} \quad v_H(\mathbf{r}) = \sum_{\mathbf{K} \neq 0} v_{\mathbf{K}} \exp(i\mathbf{K} \cdot \mathbf{r}) \quad (22)$$

The coefficients,  $\rho_{\mathbf{K}}$ , are obtained immediately from the knowledge of the occupied orbitals:

$$\psi_i(\mathbf{r}; \kappa) = \sum_{\mathbf{K}} c_{i,\mathbf{K},\kappa} \phi_{\mathbf{K}}(\mathbf{r}; \kappa) \rightarrow \rho_{\mathbf{K}} = 2 \sum_{\kappa} \sum_i |c_{i,\mathbf{K},\kappa}|^2 \quad (23)$$

and as a result of Poisson's equation:

$$\nabla^2 v_H(\mathbf{r}) = 4\pi \rho'(\mathbf{r}) \quad (24)$$

we have :

$$v_{\mathbf{K}} = -4\pi \rho_{\mathbf{K}} / K^2 \quad (25)$$

### 3.3. Gaussian functions

Two kinds of basis functions have been traditionally used in molecular calculations: Slater-type orbitals (STOs) and GTOs. They are both the product of spherical harmonics  $Y_{\ell m}(\theta, \phi)$  by a radial function  $R(r)$ . For STOs,  $R(r) = r^n \exp(-\zeta r)$ , while for the basic form of GTOs (Gaussian *primitives*),  $R(r) = r^\ell \exp(-\alpha r^2)$ . In spite of the fact that STOs reproduce much better the cusps of the wave-function in the proximity of the nuclei [it may be proven that in the vicinity of a point charge  $Z$ , the radial part of the one-electron ground state wave-function has the form:  $A(1 - Zr)$ ], their use has become less and less frequent in favour of GTOs, for which the calculation of multi-center two-electron integrals is essentially simpler. A few important characteristics of GTOs can be mentioned:

- In molecular applications, GTOs are usually constructed as a linear combination of Gaussian primitives  $g_j(\mathbf{r})$ , characterized by the same center (usually coinciding with a nucleus), the same angular numbers but different exponents:

$$\varphi_{\mu}(\mathbf{r}) = \sum_{j=1}^{p_{\mu}} d_j g_j(\mathbf{r}) ; g_j(\mathbf{r}) \equiv g(\mathbf{r}; \alpha, \ell, m) = r^{\ell} Y_{\ell m}(\theta, \phi) \exp(-\alpha_j r^2) \quad (26)$$

These combinations are often called *contracted Gaussians*, where  $\alpha_j$  represents *exponents* and  $d_j$ , *contraction coefficients*. By a suitable choice of  $d_j$  and  $\alpha_j$ , one can prepare basis functions which have, from the start, good characteristics (for example, they exhibit an almost correct cusp behavior at the core). Different conventions are used to denote the basis set. A common one can be exemplified with reference to a calculation by Rauk *et al.* concerning  $\text{NH}_3$  [39]:  $(13s8p2d/8s2p) \rightarrow [8s5p2d/4s1p]$ . The information in the round parentheses concerns primitives: for nitrogen, 13 of them are of type  $s$  ( $\ell = 0$ ), 8 of type  $p$  ( $\ell = 1, m = -1, 0, 1$ ), 2 of type  $d$  ( $\ell = 2, m = -2, -1, 0, 1, 2$ ); for hydrogen, 8 of type  $s$ , 2 of type  $p$ . The contracted set is reported in square parentheses; for hydrogen, the two  $p$  primitives have been contracted to one GTO, the eight  $s$  primitives participate in four GTOs; note that the two  $d$  primitives on nitrogen have remained uncontracted.

- In solid state applications, a finite number ( $p$ ) of GTOs are attributed to the various atoms in the reference zero cell ( $\mathbf{A}_{\mu}$  will denote the coordinate of the nucleus on which  $\varphi_{\mu}$  is centered); the same GTOs are then associated with all translationally equivalent atoms in the crystal. In total, we have  $Np$  GTOs, from which we can construct  $Np$  Gaussian-type Bloch orbitals:

$$\phi_{\mu}(\mathbf{r}; \kappa) = \sum_{\mathbf{T}} \varphi_{\mu}(\mathbf{r} - \mathbf{A}_{\mu} - \mathbf{T}) \exp(i\mathbf{\kappa} \cdot \mathbf{T}) \quad (\mu = 1, \dots, p; \kappa = 1, \dots, N) \quad (27)$$

- From experience with molecular calculations, a large number of proposals have emerged to help computational chemists in the preparation of efficient basis sets; excellent reviews exist of this subject [40, 41]. Huzinaga's sets may be mentioned [42] as an example: for each atom, several options are proposed according to the required accuracy, differing in the number of GTOs, contraction coefficients and exponents. Molecular basis sets are not directly transferable to crystalline applications. In particular, there are at least three reasons for avoiding the use of very diffuse primitives (low exponents) in solid state studies: first, the number of integrals to be explicitly calculated increases explosively; secondly, the accuracy of the calculation must be particularly high in order to avoid pseudo-linear dependence catastrophes; thirdly, diffuse functions are not of much use in densely packed crystals, because their tails are found in regions where there is large variational freedom associated with functions on other atoms. These problems are re-considered in Chapters 8 and 11.
- With respect to PWs, the use of suitably contracted GTOs permits us to describe accurately electronic distributions both in the valence and in the core region with a limited number of basis functions. The price is the loss of orthogonality, of universality and the need for more sophisticated algorithms for the calculation of the required integrals. The latter are expressed in terms

of the primitives; their evaluation is made simpler by the fundamental property that the product of two Gaussians is a Gaussian (Boys' theorem). In particular for  $s$ -type primitives:

$$g(\mathbf{r} - \mathbf{A}; \alpha, 0, 0) g'(\mathbf{r} - \mathbf{A}'; \alpha', 0, 0) = K g''(\mathbf{r} - \mathbf{A}''; \alpha'', 0, 0) \quad (28)$$

with:  $\alpha'' = \alpha + \alpha'$ ;  $\mathbf{A}'' = (\alpha \mathbf{A} + \alpha' \mathbf{A}') / \alpha''$ ; and  $K = \exp(-\alpha \alpha' |\mathbf{A} - \mathbf{A}'|^2 / \alpha'')$ . This result is easily generalized to higher quantum numbers.

- The most complicated integrals that enter the calculations are the so-called *two-electron, four-center integrals* which result from calculating the Coulomb or exchange term in the Fock Hamiltonian (equation 16) or the corresponding energy term (equation 17), after having expressed the spin orbitals  $\psi_i(\mathbf{r})$  in the GTO basis set:

$$(\mu\nu|\kappa\lambda) = \int d\mathbf{r} d\mathbf{r}' \{ [\varphi_\mu(\mathbf{r} - \mathbf{A}_\mu - \mathbf{T}_\mu) \varphi_\nu(\mathbf{r} - \mathbf{A}_\nu - \mathbf{T}_\nu)] \times \\ \times [\varphi_\kappa(\mathbf{r}' - \mathbf{A}_\kappa - \mathbf{T}_\kappa) \varphi_\lambda(\mathbf{r}' - \mathbf{A}_\lambda - \mathbf{T}_\lambda)] / |\mathbf{r} - \mathbf{r}'| \} \quad (29)$$

This kind of integral, as well as the simpler nuclear attraction ones, are obtained by starting from an auxiliary function of one variable:

$$F_0(w) = \int_0^1 ds \exp(-ws^2) \quad (30)$$

with the help of powerful recurrence relations [43]. A large part of the success of GTOs in molecular and solid state QM applications is due precisely to the efficiency of the algorithms developed for integral calculation.

- An additional advantage of GTOs is due to the fact that their Fourier transform is another Gaussian [ $F_p \{ \exp(-\alpha r^2) \} \propto \exp(-p^2/4\alpha)$ ], and their use in combination with PW techniques is therefore easy [29].
- A given set of GTOs, associated with atomic positions, usually performs better when the atoms are close to each other than if they are far apart. This is a consequence of basis set incompleteness; the region between two atoms is better described if use is made of functions centered on both atoms. There is then an over-estimation of binding energies, which is called basis-set superposition error (BSSE). BSSE can be very important with poor basis sets; the counterpoise technique for estimating this error gives usually reasonable results [41, 44].

## 4. Standard Methods of Solution

### 4.1. $\kappa$ -factorization of the one-electron equations for crystals

Although DF and HF schemes correspond to totally different approaches from a theoretical point of view, in practical applications there are profound similarities between them. They are both specially designed for the GSES, and have little significance for excited states; in both cases, the one-electron hamiltonian depends functionally on the GSES, and the equations must be solved through an SCF procedure; the basis set problems are similar. We describe in this section standard methods of solution for periodic systems in very general terms, without

specifying which one-electron Hamiltonian  $\hat{h}$  we are referring to, except where needed. A more detailed description is provided in Chapters 8, 9 and 10.

As shown in Chapter 2, BFs associated with different  $\kappa$  points within the first BZ belong to different irreducible representations of the group of one-electron Hamiltonians,  $\hat{h}$ . It is then possible to factorize the problem into separate parts for each  $\kappa$ :

- Consider the  $p$  BFs,  $\phi_\mu(\mathbf{r}; \kappa)$ , associated with  $\kappa$  (equation 21 or 23);
- Calculate the matrix elements:  $H_{\mu\nu}^\kappa = \langle \phi_\mu^\kappa | \hat{h} | \phi_\nu^\kappa \rangle$  ;  $S_{\mu\nu}^\kappa = \langle \phi_\mu^\kappa | \phi_\nu^\kappa \rangle$  (the latter equals  $\delta_{\mu,\nu}$  in the PW set);
- Solve the  $p \times p$  matrix equation:

$$H^\kappa C^\kappa = S^\kappa C^\kappa E^\kappa, \quad (31)$$

where the diagonal matrix  $E^\kappa$  contains the eigenvalues  $\varepsilon_i^\kappa$ , and the matrix  $C^\kappa$  contains, columnwise, the coefficients of the crystalline orbitals (COs):

$$\psi_i(\mathbf{r}; \kappa) = \sum_{\mu=1}^p c_{\mu i}^\kappa \phi_\mu(\mathbf{r}; \kappa). \quad (32)$$

The above procedure should be carried on for the complete set of  $\kappa$  points in the first BZ, so as to determine the complete set of COs (that is, the KS or HF spin-orbitals: see equations 11 and 16) with the precision granted by the basis set adopted and by the accuracy of the algorithms. There are, however, some points which need closer analysis:

- The problem of how many and which  $\kappa$  points should be considered is an extremely important one. In order to reconstruct the  $\hat{h}$  operator and to calculate the crystalline properties from the solution, we need all occupied spin-orbitals, in principle. Due to the continuity of eigenvalues and eigenfunctions with respect to  $\kappa$  (see Chapter 2), it is, however, possible to derive the required information from the results obtained at a few suitably sampled  $\kappa$  points, as discussed in the next subsection.
- In order to calculate  $H_{\mu\nu}^\kappa$  matrix elements, we must consider functions and operators which are extended to the whole crystalline structure. Except in particularly favorable cases, we are faced with a problem of summations over one or more indices associated with the different crystalline cells. The accurate and efficient handling of such series (*lattice sums*) often determines the final quality of a computational scheme (section 4.3).
- It has been mentioned already that the process must be repeated until self-consistency is achieved, that is, until eigenvalues and eigenvectors coincide, within a given tolerance, with those used for the reconstruction of  $\hat{h}$ . This may be very difficult to achieve at times. All techniques that accelerate convergence are valuable (section 4.4).
- To solve the matrix equation 31, the diagonalization of  $p \times p$  complex matrices is needed. The time required is proportional to  $p^3$ . With very complex systems,  $p$  can become very large, and the diagonalization step will be the slowest one, perhaps to the point where the system is untreatable. Different, non-standard procedures, which by-pass this step, may then become preferable (section 5).

#### 4.2. Reciprocal space integration and interpolation

From the knowledge of the eigenvalues,  $\varepsilon_i^{\mathbf{k}}$  and the eigenfunctions,  $\psi_i(\mathbf{r}; \mathbf{k})$ , at a few, sample  $\mathbf{k}$  points (to be indicated in the following by  $\underline{\mathbf{k}}$ ), we want to obtain accurate estimates of quantities such as the number of states below a certain energy,  $i(e)$  and the matrix of integrated densities of states (IDOS),  $I(e)$ , which imply a sum over *all*  $\mathbf{k}$  points:

$$\begin{aligned} i(e) &= 2V_{BZ}^{-1} \sum_i \int_{BZ} d\mathbf{k} \theta(e - \varepsilon_i^{\mathbf{k}}) \\ I_{\mu\nu}^T(e) &= 2V_{BZ}^{-1} \sum_i \int_{BZ} d\mathbf{k} c_{\mu i}^{\mathbf{k}*} c_{\nu i}^{\mathbf{k}} \exp(i\mathbf{k} \cdot \mathbf{T}) \theta(e - \varepsilon_i^{\mathbf{k}}) \end{aligned} \quad (33)$$

The sum has been replaced with an integral over the  $BZ$ , due to the fact that  $\mathbf{k}$  points are uniformly distributed in reciprocal space (see Chapter 2); the condition that only the orbitals of energy less than  $e$  are included in the sum is expressed by the presence in the integrand of the step function  $\theta(e - \varepsilon_i^{\mathbf{k}})$ , whose value is 1 if  $\varepsilon_i^{\mathbf{k}}$  is less than  $e$ , and is zero otherwise. The quantities  $i(e)$  and  $I(e)$  are very important. The Fermi energy,  $e_F$ , is determined by imposing the condition:  $i(e_F) = 2n$ , that is, by requiring that there are exactly  $2n$  spin-orbitals per cell with energy less than  $e_F$ . As shown in Chapter 2, the total density of states (DOS):  $n(e) = di(e)/de$  and the IDOS derivative,  $N(e) = dI(e)/de$ , which is called the projected density of states (PDOS), give rich information on the chemical structure of the system and allow all one-electron properties to be obtained within the independent-electron approximation [27]. The value of  $I(e)$  at the Fermi energy,  $e_F$ , is the  $P$  matrix [ $P = I(e_F)$ ]. When using a basis set of localized functions, the one-matrix,  $\gamma_0(\mathbf{r}, \mathbf{r}')$ , (equation 3) corresponding to a detor constructed on the occupied HF or KS spin-orbitals, is conveniently expressed in terms of the  $P$  matrix:

$$\gamma_0(\mathbf{r}, \mathbf{r}') = \sum_{\mu, \nu=1}^p \sum_{\mathbf{T}, \mathbf{T}'} P_{\mu\nu}^{T'-T} \varphi_{\mu}(\mathbf{r} - \mathbf{A}_{\mu} - \mathbf{T}) \varphi_{\nu}(\mathbf{r}' - \mathbf{A}_{\nu} - \mathbf{T}')$$

The determination of the Fermi energy,  $e_F$ , is a delicate problem only in the case of metals (see Chapter 4 for details). For insulators,  $n$  bands are fully occupied, the others are void. The  $\theta(e - \varepsilon_i^{\mathbf{k}})$  function can then be dropped from the integrand and the sum over  $i$  is limited to the  $n$  lowest eigenvalues at each  $\mathbf{k}$  point.

In summary, we must estimate integrals of the form:

$$\Xi(e) = V_{BZ}^{-1} \int_{BZ} d\mathbf{k} [\xi(\mathbf{k}) \theta(e - \varepsilon(\mathbf{k}))] = V_{BZ}^{-1} \int_{BZ'(e)} d\mathbf{k} \xi(\mathbf{k}) \quad (34)$$

where  $\xi(\mathbf{k})$  and  $\varepsilon(\mathbf{k})$  are well-behaved, periodic functions in reciprocal space, and  $BZ'(e)$  is the portion of  $BZ$  where  $\varepsilon(\mathbf{k}) < e$ : it coincides with the whole  $BZ$ , if  $\varepsilon(\mathbf{k})$  is less than  $e$  everywhere and vanishes if  $\varepsilon(\mathbf{k})$  is always greater than  $e$ .

If  $BZ'(e) = BZ$ , *special points* techniques may be used [45]. This means that one may select a special set of points  $\{\underline{\mathbf{k}}\}$  within the  $BZ$ , with a weight,  $w(\underline{\mathbf{k}})$ ,

associated with each of them, evaluate  $\xi(\underline{\mathbf{k}})$  at each of them and substitute the integral with a weighted sum:  $\Xi \approx \sum_{\underline{\mathbf{k}}} w(\underline{\mathbf{k}}) \xi(\underline{\mathbf{k}})$ . A discussion of these techniques and a documentation of how good their performance can be is provided in Chapter 4.

In the more general case ( $BZ'(e) \neq BZ$ ), the linear [46] or quadratic [47] tetrahedron techniques are usually adopted. The  $BZ$  is subdivided into tetrahedral mini-cells: the integral will be the sum of sub-integrals over each of them. After evaluating  $\xi$  and  $\varepsilon$  at the vertices of the tetrahedra,  $\underline{\mathbf{k}}$ , a linear or quadratic approximation is obtained for both  $\xi(\mathbf{k})$  and  $\varepsilon(\mathbf{k})$  inside the tetrahedron: the sub-integral is finally evaluated analytically using these approximate expressions. An alternative method can be used for evaluating directly  $\Xi$  integrals, starting from the value of  $\xi(\mathbf{k})$  and  $\varepsilon(\mathbf{k})$  at a few special points,  $\underline{\mathbf{k}}$ : this is called the Fourier-Legendre expansion technique, because it is based first, on a Fourier representation of bands, second, on the expansion of the DOS associated with each band into an orthonormal set of Legendre functions [48, 27]. It is not as accurate as the tetrahedron technique, but has the advantage of providing analytic expressions for the  $\Xi$  dependence on  $e$  and of being much faster.

#### 4.3. Lattice sums

The problem of Coulomb lattice sums is present in both DF and HF schemes and must be treated with extreme care, because the Coulomb series is only conditionally convergent. In HF theory, we have the additional problem of the exchange series, which is involved in the construction of the Fock hamiltonian,  $\hat{f}$  (see equation 16), and gives a contribution,  $E_{ex}$ , to the HF energy (see the last term in equation 17). Since both these problems are discussed explicitly in Chapter 8, we will not examine them here.

#### 4.4. The SCF problem

In variational approaches like KS or HF, the SCF procedure can be viewed as a path in the space of the variational coefficients  $\{c_{\mu i}\} \equiv \mathbf{c}$  (for simplicity, the  $\mathbf{k}$  index is dropped). This path begins at some initial good (?) guess ( $\mathbf{c} = \mathbf{c}_{\text{start}}$ ) and ends, hopefully, at  $\mathbf{c} = \mathbf{c}_{\text{scf}}$ , where the minimum of the functional  $E_0[\mathbf{c}]$  occurs, subject to the orthonormality condition of the spin-orbitals:

$$\sum_{\mu, \nu=1}^p c_{\mu i}^* c_{\nu j} S_{\mu\nu} = \delta_{ij}$$

At  $\mathbf{c}_{\text{scf}}$ , the self-consistency condition is satisfied. On the other hand, self-consistency does not necessarily imply that a minimum has been reached, although this is always the case, in practice. The standard movement of the representative point  $\mathbf{c}$  in the coefficient space is as described in section 4.1. At the general  $n$ -th cycle of the SCF procedure, the current point  $\mathbf{c}_n$  is used to generate the new hamiltonian,  $\hat{h}_n$ ; this in turn (equation 25) generates  $\mathbf{c}_{n+1}$ , satisfying the orthonormality constraint. There is, however, no guarantee that the new point moves in the direction of  $\mathbf{c}_{\text{scf}}$ . In fact, critical cases are frequent: oscillations in

energy occur, requiring a large number of steps in the procedure, sometimes of the order of hundreds or even ending in a numerical catastrophe. Many procedures have been devised to accelerate the SCF procedure. Some of them simply damp oscillations by mixing quantities belonging to the current and preceding cycles before starting the new one (typically, the Hamiltonian matrix in HF calculation and the density in DF calculations):  $\hat{h}_n \leftarrow \alpha \hat{h}_{n-1} + (1-\alpha) \hat{h}_n$ , where the mixing coefficient  $\alpha$  may be larger than 0.5. More sophisticated procedures have been proposed in the field of molecular calculations, which try to obtain good guesses for  $c$  for use in subsequent cycles by extrapolation [49]. These techniques are just now beginning to be used in solid state applications. In the next section, we present briefly the Car-Parrinello technique, which combines the SCF problem with the problem of the optimization of nuclear coordinates in an original way.

## 5. Non-standard Methods of Solution

To give an idea of the variety of approaches that can be used for the *ab initio* treatment of periodic systems, a few of those in current practice are sketched here.

### KKR (Korringa-Kohn-Rostoker)

There are some aspects of this powerful technique [50] which make it awkward to use for those accustomed to the ordinary tools and concepts of molecular quantum chemistry (for simplicity, we consider its application to elementary crystals with one atom per cell, in the framework of DF theory). First, space is partitioned into *atomic spheres* centered on atoms, and a complementary *interstitial region*; the atomic sphere is the largest that can be inscribed in the Wigner-Seitz cell. Secondly, a simplified form is assumed for  $V_{eff}(\mathbf{r})$  (equation 13), that is,  $V_{eff}(\mathbf{r})$  is zero in the interstitial region and is a radial function [ $V_{eff}(\mathbf{r}) = V(r_A)$ ] inside the atomic spheres: this is called the muffin-tin (MT) approximation. Thirdly, for the solution of the KS equations, Green-function techniques are adopted, that is, the eigenfunction at a given  $\kappa$  point, corresponding to an eigenvalue  $\epsilon < 0$ ,  $\psi(\kappa, \mathbf{r}; \epsilon)$ , is the solution of the Lippmann-Schwinger integral equation:

$$\psi(\kappa, \mathbf{r}; \epsilon) = \int d\mathbf{r}' \{ [-\exp(-|\mathbf{r} - \mathbf{r}'| \sqrt{-\epsilon}) / (4\pi|\mathbf{r} - \mathbf{r}'|)] V_{eff}(\mathbf{r}') \psi(\kappa, \mathbf{r}'; \epsilon) \} \quad (35)$$

The expression in square brackets is the Green function for the free electron. The MT form of the potential and the fact that:

$$\psi(\kappa, \mathbf{r} + \mathbf{T}; \epsilon) = \exp(i\kappa \cdot \mathbf{T}) \psi(\kappa, \mathbf{r}; \epsilon)$$

for all lattice vectors,  $\mathbf{T}$ , as a result of Bloch's theorem (see Chapter 2, equation 41), are now taken into account. The Lippmann-Schwinger equation can then be recast in a very convenient form, where the integral is restricted to the interior of a single atomic sphere, in the reference zero cell. After expressing  $\psi(\kappa, \mathbf{r}; \epsilon)$  as a linear combination of "partial waves", the eigenvalues,  $\epsilon$ , are obtained by looking for non-trivial solutions, which correspond to the zeros of a determinant ( $|I - G(\epsilon)V| = 0$ : see Chapter 13, Appendix 1). A remarkable example of the

power of this technique is provided by the detailed study of the electronic properties of 32 elementary metals from hydrogen to indium, by Moruzzi, Janak and Williams [51].

### APW (augmented plane waves)

The partition of space between atomic spheres and the interstitial region is the characteristic feature of a whole family of techniques, derived from Slater's original proposal [52] through continuous improvements and transformations: APW, LAPW (linearized plane waves) [53], FLAPW (Full potential LAPW) [54]. Essentially, a plane-wave expansion is adopted in the interstitial region, while spherical harmonics multiplied by radial functions are used within the spheres; the continuity of the wave-function and its derivative across the boundary is the principal condition to be obeyed. All restrictions concerning the form of the potential have been progressively removed, complicated crystalline structures can be treated without problems and linearized forms of the equations have been adopted which allow all eigenvalues at a given  $\kappa$  to be found at the same time through matrix diagonalization. These techniques are specially suited to the study of properties related to the distribution of inner electrons, such as the electric field gradient at the nuclear positions [55]. The FLAPW method is discussed in detail in Chapter 9.

### Recursion

This method shares with KKR the use of Green function techniques. Consider a system described by one-electron equations, and a representative basis set of localized functions  $\phi_\mu(\mathbf{r})$ . The originality of the recursion technique [56] lies in the fact that its aim is not the determination of eigenvalues and eigenfunctions, but rather, to determine the PDOS directly (section 4.2):

$$N_{\mu\nu}(e) = \sum_i c_{\mu i}^* c_{\nu i} \delta(e - \epsilon_i)$$

For each  $e$ , the sparse nature of the  $H$  and  $S$  matrices is exploited. Starting from a given "centre",  $\phi_\mu$ , one considers its neighbours (those that have non-negligible matrix elements with  $\phi_\mu$ ), then the next nearest neighbours, and so on. Through a continued fraction expansion, the PDOS is determined with progressively better precision and the procedure is truncated after considering a certain number of steps. An appealing aspect of this method is that it does not exploit translational symmetry, except for the determination of  $H$  and  $S$  matrices, and is largely insensitive to boundary conditions: for these reasons, it can be applied quite naturally to the study of defects in solids. A drawback lies in the requirement of matrix sparseness, which has made difficult its use with *ab initio* Hamiltonians, until very recently [57].

### CP (Car-Parrinello)

The CP technique [58] differs from those described up to now because it treats the problem of the nuclear motion and that of the electron response to nuclear displacements contemporaneously. It is expedient to give a short account of this method which is gaining enormous importance in solid state physics; for more details, reference can be made to the review article by Remler and Madden [59]. As anticipated in section 4.4, the SCF process, which consists of finding the optima

$c_i$  coefficients of the basis functions defining the occupied COs, can be viewed as a motion of the system in the coefficient space towards the energy minimum, subject to an orthonormality constraint. Such problems of constrained motion can be treated in a classical Lagrangian formalism, by introducing a fictitious time,  $t$ , a fictitious mass,  $\mu$  and a fictitious kinetic energy,  $\mu \cdot \sum_i (\dot{c}_i)^2/2$ . At each time step,  $dt$ , the electronic variables evolve in a direction dictated by the acting forces ( $F_i = -\partial E[c]/\partial c_i$ ) which drive the system towards lower total energy, subject to constraints and to *inertial forces*. *Attrition* terms are included in the Lagrangian, so that the system evolves towards an energy minimum, hopefully the absolute one. With fixed nuclei, this procedure is usually much less efficient than the standard one, described in section 4.4.

The brilliant idea in the CP method was to include in the Lagrangian both the real nuclear motion in R-space and the fictitious electron motion in c-space, on the same footing. The system (nuclei + electrons) starts from a given initial configuration ( $R, \dot{R}; c, \dot{c}$ ) and evolves in the phase space: the forces acting on nuclei are calculated using the Hellman-Feynman theorem [59]. This process is also called *simulated annealing*, because the initial configuration can be made to correspond to a given temperature, and the attrition terms are, therefore, “cooling” the system by subtracting real and fictitious kinetic energy. With a proper choice of the time step  $dt$ , of the fictitious mass  $\mu$ , and of the attrition terms, the nuclei are found to move close to the ever-changing Born-Oppenheimer energy surface,  $E_0(R)$  (see equation 2) and eventually to reach a local energy minimum. At a stroke, both the electronic SCF problem has been solved and the equilibrium configuration has been reached, an achievement that could require painstaking calculations with standard techniques. A fringe benefit is that one obtains the electron response to nuclear motion automatically.

The large majority of applications of the CP method are concerned with this energy minimization problem. However, the same technique can be used for studying real problems of molecular dynamics, such as the distribution of phonon frequencies (see for instance reference [60] concerning highly compressed water ice). In solid state applications, the standard ingredients of the CP method are:

- 1) the DF-LDA Hamiltonian (derivatives are easily calculated);
- 2) the super-cell technique (in the different super-cells, nuclei and electrons are moving identically; only, the  $\kappa = 0$  point is considered in integrations);
- 3) the soft-core pseudo-potentials for core electrons and plane waves for valence electrons (these choices permit the use of a simple, orthonormal, universal basis set, and of highly efficient fast-Fourier-transform algorithms).

### GFMC and VMC (Green-Function Montecarlo and Variational Montecarlo)

The impact of Montecarlo techniques in solid state physics is mostly due to the calculations by Ceperley and Alder, which have provided information of unprecedented accuracy concerning the homogeneous and inhomogeneous electron gas at low densities [11] (see section 2.1). These techniques allow the  $2n$ -electron closed-shell wavefunction,  $\Psi_0$ , to be treated without the need for separation into the products of one-electron wave-functions. This is possible because integrals of the type  $\langle \Psi_0 | \hat{H}^{nr} | \Psi_0 \rangle$  in  $3n$  spatial co-ordinates are evaluated accurately and economically by a statistical sampling procedure, the so-called Metropolis algorithm. There are two fundamental types of Montecarlo methods. According to GFMC, the evolution of the wave-function in imaginary time,  $\tau = it$ , is followed using statistical techniques, starting from a trial function,  $\Psi_T$ . For  $\tau \rightarrow \infty$ , all

components in the wave-function, except  $\Psi_0$ , are damped because of the time evolution operator:  $\exp(-iHt) = \exp(-H\tau)$ .

This procedure has been adopted, for instance, in the electron-gas calculations just mentioned. In VMC, the trial wave-function is made dependent on some variational parameter,  $q$  [ $\Psi_T \equiv \Psi(q)$ ]. The optimal value of the parameters is determined by minimizing the expectation value of the energy:

$$E(q) = \langle \Psi(q) | \hat{H}^{nr} | \Psi(q) \rangle / \langle \Psi(q) | \Psi(q) \rangle$$

The trial function is usually written so as to include explicitly the most important part of the correlation. The form most often adopted is the product of a Jastrow factor,  $J(q)$  [61], by a single determinantal approximate wave-function, usually the HF solution, perhaps with the introduction of adjustable scaling factors,  $q'$ :  $\Psi_T \equiv J(q)\Psi_{HF}(q')$ . The Jastrow factor is totally symmetric in the electron coordinates (therefore,  $\Psi_T$  is antisymmetric) and has the form  $\sum_{i>j} [ar_{ij}/(1+br_{ij})]$ , where  $a$  and  $b$  are variational parameters. Thanks to this factor, the trial wave-function automatically contains electron-electron *correlation cusps* of the correct form, that is, each electron moves accompanied by a correlation hole. The main difficulty in the application of Montecarlo techniques to solid state problems, is that they are very inefficient in the case of heavy atoms where a large majority of the sampling points in the statistical integration would end in core regions, the most uninteresting ones: computational requirements increase very rapidly with  $Z$  ( $\propto Z^{5.5}$ ). The use of pseudo-potentials would allow us to get rid of core-electrons, but their non-local form (equation 19) makes their use with standard Montecarlo techniques very difficult. Nevertheless, interesting attempts have been made in this direction: pseudo-potential calculations of crystalline silicon have been performed both using VMC [62], and GFMC techniques [63].

### N-scaling

When one considers such a bewildering variety of techniques, the question can be raised as to which one has best performance or if there is a theoretical limit to the efficiency of new techniques that can be devised. In a system containing  $N$  non-equivalent atoms, all physically relevant information is proportional to  $N$  (in the rest of this lecture,  $N$  designates the number of crystalline cells; here we use it in a different sense, to conform with standard nomenclature). However, standard schemes depend on  $N^3$ , due to the necessity of determining a number of orbitals proportional to  $N$ , which involves diagonalization or inversion of  $(kN) \times (kN)$  matrices (equation 31: the pre-factor  $k$  depends on the basis set size). Much work has taken place in recent years in the framework of DF-LDA approaches, aimed at devising algorithms which vary linearly with  $N$ , at least for very large  $N$  values. There is a variety of proposals, still at a rough stage of implementation [64]. When such tools become available, it will be possible to study very complex systems which are beyond the accessibility of *ab initio* schemes at present.

## 6. From the Electronic Structure to the Crystalline Properties

A large part of this book is devoted to the calculation of ground state crystalline properties: here, we make a few general remarks. Consider an observable quantity

$O$ , depending on the electronic variables. Its average ground state value,  $O^{av}$ , is given by the expression:

$$O^{av} = \langle \Psi_0 | \hat{O} | \Psi_0 \rangle \quad (36)$$

where the operator  $\hat{O}$  is obtained from the classical expression of the observable  $O$ , via the correspondence principle:  $\mathbf{r} \rightarrow \mathbf{r}$ ;  $\mathbf{p} \rightarrow i\nabla$  (for instance:  $T = \sum_i |\mathbf{p}_i|^2/2 \rightarrow -\sum_i \nabla_i^2/2$ ). If  $O$  is a one-electron observable, that is,  $O \equiv O^1 = \sum_i o_i$ , which is often the case, then expression 36 can be simplified by making reference to the one-electron density matrix,  $\gamma_0(\mathbf{r}, \mathbf{r}')$  (equation 3):

$$O^{1av} = \int d\mathbf{r} [\hat{o}(\mathbf{r}') \gamma_0(\mathbf{r}, \mathbf{r}')]_{\mathbf{r}' \rightarrow \mathbf{r}} \quad (37)$$

In HF theory,  $\Psi_{HF}$  is an approximation of  $\Psi_0$  in the form of an antisymmetrized product of spin-orbitals. It is easily seen that:

$$\gamma_{HF}(\mathbf{r}, \mathbf{r}') = 2 \sum_i \psi_{HF,i}^*(\mathbf{r}) \psi_{HF,i}(\mathbf{r}')$$

for closed-shell systems, the sum being restricted to the  $n/2$  occupied orbitals and we can write, for one-electron operators:

$$O_{HF}^{1av} = 2 \sum_i \int d\mathbf{r} \psi_{HF,i}^*(\mathbf{r}) \hat{o}(\mathbf{r}) \psi_{HF,i}(\mathbf{r}) \quad (38)$$

A similar expression is obtained for two-electron observables.  $O_{HF}^{1av}$  will be affected by a correlation error which can be estimated according to the usual techniques.

When considering DF theory, Hohenberg and Kohn's first theorem (section 2.2) insures that for each observable, a functional  $O([\rho_0]; \mathbf{r})$  must exist such that:

$$O^{av} = \int d\mathbf{r} O([\rho_0]; \mathbf{r}) \rho_0(\mathbf{r}) \quad (39)$$

However, there is no general and simple prescription for generating such functionals. By considering the formal similarity of HF and KS procedures, it would be tempting to use equation 38 again, this time making reference to the KS pseudo-orbitals  $\psi_{KS,i}(\mathbf{r})$ . This seems to be impossible because  $\Psi_{KS}$  is *not* an approximation of  $\Psi_0$ . However, it turns out that it is, in fact, convenient to isolate such a contribution, and to write [65]:

$$O^{1av} = 2 \sum_i \int d\mathbf{r} \psi_{KS,i}^*(\mathbf{r}) \hat{o}(\mathbf{r}) \psi_{KS,i}(\mathbf{r}) + \Delta O[\rho] \quad (40)$$

Different prescriptions are proposed to estimate the correction  $\Delta O$ ; in LDA:

$$\Delta O_{LDA}[\rho] = \int d\mathbf{r} [o_h(\rho(\mathbf{r})) - o_f(\rho(\mathbf{r}))] \quad (41)$$

where  $o_h(\rho)$  and  $o_f(\rho)$  are the average value (per electron) of the observable  $O$ , calculated for the homogeneous gas of interacting electrons and of free electrons, respectively. Similar expressions, with the addition of a SIC term (see section

2.2), have been used, for example, to calculate Compton profiles from KS pseudo-wave-functions (see Chapter 13).

There are important quantities, notably the macroscopic dielectric polarization, which cannot be expressed in the form of equation 36: in these cases, the availability of an approximate form of the wave-function, as provided by HF theory, allows a more satisfactory formulation of the theory (see Chapter 14).

## 7. Conclusions

There has been, in recent years, an explosive expansion in the field of *ab initio* techniques for the treatment of solid state problems. Very rapidly, new proposals emerge, existing schemes become obsolete, or conversely, receive new impetus from the efficient solution of a certain critical computational step. The solution of problems which appeared untreatable just a few years ago is now commonplace. For users not directly involved in the development of the methodologies and in the implementation of computer codes, such a rapid evolution may be disturbing. Getting accustomed to the use of one of these complex programs, knowing how to obtain answers to one's problems and how to interpret results involves an important human investment in terms of study, time and practice. The simplicity of input, the transparency of output and the quality of the documentation can be more important than a factor two, say, in speed.

It is fortunate that both in the field of DF and of HF approaches, powerful, general-purpose programs are now available, which tend to incorporate in, consecutive releases, the natural evolution of techniques in the field. The "life-time" of these projects is becoming much longer (of the order of ten years at least) and justifies the effort required to become familiar with the programs. Another advantage of these public codes is that one can compare their performance over a variety of systems. Due to their *ab initio* nature, they should give comparable results when applied to the same system. There is now the trend to test these powerful codes with whole families of crystalline compounds. We may cite a few examples from recent literature: seventeen III-V and IV-IV semiconductors (LCAO-HF) [66]; nine oxides (LCAO-HF) [67]; ten wurtzite crystals (LCAO-LDA) [68]; carbides and nitrides of eight transition metals [69]; all hcp metals from Be to Cd [55]. This kind of documentation may help final users to make an unbiased, sensible choice. Having at least two codes based on different schemes available would be preferable, of course, because the optimal area of application is usually different for the different approaches.

## Acknowledgments

This work was supported in part by the Italian MURST (fondi 40%), and by the Italian CNR.

## References

1. A. Szabó and N.S. Ostlund, *Modern Quantum Chemistry*, McGraw-Hill, New York (1993).
2. R.E. Moss, *Advanced Molecular Quantum Mechanics*, Chapman and Hall, London (1973).

3. P.P. Singh, Phys. Rev. B **49**, 4954 (1994).
4. J.C. Inkson, *Many-body Theory of Solids*, Plenum, New York (1984).
5. L. Hedin, Phys. Rev. **139**, A796 (1965).
6. P. Fulde, *Electron Correlations in Molecules and Solids*, Solid-State Sciences, Vol. 100, Springer, Berlin (1993).
7. S. Raimès, *The Wave Mechanics of Electrons in Metals*, North Holland, Amsterdam (1961).
8. H.J. Monkhorst, Phys. Rev. B **20**, 1504 (1979).
9. J. Delhalle and J.L. Calais, J. Chem. Phys. **85**, 5286 (1986).
10. M. Gell-Mann and K.A. Bruckner, Phys. Rev. **106**, 364 (1957).
11. D.M. Ceperley, Phys. Rev. B **18**, 3126 (1978); D.M. Ceperley and B.J. Alder, Phys. Rev. Lett. **45**, 566 (1980).
12. J.P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).
13. P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
14. W. Kohn and L.J. Sham, Phys. Rev. **140**, A1133 (1965).
15. L.H. Thomas, Proc. Cambr. Phil. Soc. **23**, 542 (1927); E. Fermi, Z. f. Phys. **49**, 550 (1928).
16. L. Hedin and B.I. Lundqvist, J. Phys. C: Solid St. Phys. **4**, 2064 (1971).
17. J. C. Slater, Phys. Rev. **81**, 385 (1951).
18. C. Lee, W. Yang and R.G. Parr, Phys. Rev. B **37**, 329 (1975); A.D. Becke, J. Chem. Phys. **88**, 1053 (1988); L.C. Wilson and M. Levy, Phys. Rev. B **41**, 12930 (1990); J.P. Perdew and Y. Wang, Phys. Rev. B **45**, 13244 (1991).
19. O. Gunnarsson and B.I. Lundqvist, Phys. Rev. B **13**, 4274 (1976).
20. M.R. Pederson and C.C. Lin, J. Chem. Phys. **88**, 1807 (1988); E.S. Fois, J.I. Penman and P.A. Madden, J. Chem. Phys. **98**, 6352 (1993).
21. B.G. Johnson, C.A. Gonzales, P.M.W. Gill and J.A. Pople, Chem. Phys. Lett. **221**, 100 (1994); M. Biagini, Phys. Rev. B **49**, 2156 (1994).
22. M.L. Cohen, in *Structure and Bonding in Crystals*, Vol I (M. O'Keeffe and A. Navrotsky, eds), Academic, New York, p.25 (1981).
23. J.M. André, L. Gouverneur and G. Leroy, Intern. J. Quantum Chem. **1**, 451 (1967); J.L. Brédas, J.M. André, J.G. Fripiat and J. Delhalle, Gazz. Chim. Ital. **108**, 307 (1978).
24. G. Del Re, J. Ladik and G. Biczó, Phys. Rev. **155**, 997 (1967).
25. R.N. Euwema, D.L. Wilhite and G.T. Surratt, Phys. Rev. B **7**, 818 (1973).
26. A.B. Kunz, Phys. Rev. B **6**, 606 (1972).
27. C. Pisani, R. Dovesi and C. Roetti, *Hartree-Fock ab initio Treatment of Crystalline Systems*, Lecture Notes in Chem. Vol. 48, Springer, Berlin (1988).
28. R. Dovesi, C. Pisani, C. Roetti, M. Causà and V.R. Saunders, *CRYSTAL 88, Program No. 577*, QCPE, Indiana University, Bloomington, IN (1989); R. Dovesi, C. Roetti, and V.R. Saunders, *CRYSTAL 92 User Documentation*, University of Torino, Torino (1992).
29. S. Baroni, G. Pastori Parravicini and G. Pezzica, Phys. Rev. B **32**, 4077 (1985).
30. R. Hott, Phys. Rev. B **44**, 1057 (1991).
31. P. Piecuch, S. Zarrabian, J. Paldus and J. Cizek, Phys. Rev. B **42**, 3351 (1990).
32. M.B. Lepetit, E. Aprà, J.P. Malrieu and R. Dovesi, Phys. Rev. B **46**, 12974 (1992).
33. M. Causà, R. Dovesi and C. Roetti, Phys. Rev. B **43**, 11937 (1991).
34. G.B. Bachelet, D.R. Hamann and M. Schlüter, Phys. Rev. B **26**, 4199 (1982); D.R. Hamann, Phys. Rev. B **40**, 2980 (1989).
35. D. Vanderbilt, Phys. Rev. B **41**, 7892 (1990).
36. P. Durand and J.C. Barthelat, Theor. Chim. Acta **38**, 283 (1975); J.C. Barthelat, P. Durand and A. Serafini, Molec. Phys. **33**, 159 (1977).
37. P.J. Hay and W.R. Wadt, J. Chem. Phys. **82**, 270; 284; 299 (1985).
38. G. te Velde and E.J. Baerends, Phys. Rev. B **44**, 7888 (1991).
39. A. Rauk, L.C. Allen and E. Clementi, J. Chem. Phys. **52**, 4193 (1970).
40. S.F. Boys and F. Bernardi, Molec. Phys. **19**, 553 (1970).
41. E.R. Davidson and D. Feller, Chem. Rev. **86**, 681 (1986).
42. S. Huzinaga, *Gaussian Basis Sets for Molecular Calculations*, Elsevier, Amsterdam (1984).
43. L.E. McMurchie and E.R. Davidson, J. Comp. Phys. **26**, 218 (1978); **44**, 289 (1981); V.R. Saunders, in *Methods in Computational Molecular Physics* (G.H.F. Diercksen and S. Wilson, Eds.) Reidel, Dordrecht (1983).
44. J. Tomasi, in *Basi Teoriche ed Applicazioni Numeriche* (R. Moccia and C. Pisani, Eds.) Polo Editoriale Chimico, Milano, p. 129 (1990).
45. A. Baldereschi, Phys. Rev. B **7**, 5212 (1973); D.J. Chadi and M.L. Cohen, Phys. Rev. B **16**, 1746 (1977); H.J. Monkhorst and J.D. Pack, Phys. Rev. B **13**, 5188 (1976); C. Pisani and P. Ugliengo, Phys. Stat. Sol. (b) **96**, 537 (1982).
46. O. Jepsen and O.K. Andersen, Solid St. Comm. **9**, 1763 (1971); G. Lehmann and M. Taut; Phys. Status Sol. (b) **54**, 469 (1972).
47. M.H. Boon, M.S. Methfessel and F.M. Müller, J. Phys. C **19**, 5337 (1986); G. Wiesenekker, G. te Velde and E.J. Baerends, J. Phys. C **21**, 4263 (1988).
48. G. Angonoo, R. Dovesi, C. Pisani and C. Roetti, Phil. Mag. **44**, 413 (1981).
49. I.H. Hillier and V.R. Saunders, Int. J. Quantum Chem. **4**, 503 (1970); R. Seeger and J.A. Pople, J. Chem. Phys. **65**, 265 (1976); P. Pulay, J. Comp. Chem. **3**, 556 (1982); J. Cioslowski, J. Chem. Phys. **89**, 2126 (1988); H.B. Schlegel and J.J.W. Douall, *Computational Advances in Organic Chemistry*, Kluwer, New York (1991).
50. J. Korrington, Physica **13**, 392 (1947); W. Kohn and N. Rostoker, Phys. Rev. **94**, 1111 (1954).
51. V.L. Moruzzi, J.F. Janak, A.R. Williams, *Calculated Electronic Properties of Metals*, Pergamon, New York (1978).
52. J.C. Slater, Phys. Rev. **51**, 846 (1937).
53. O.K. Andersen, Phys. Rev. B **12**, 3060 (1975).
54. E. Wimmer, H. Krakauer, M. Weinert and A.J. Freeman, Phys. Rev. B **24**, 864 (1981); P. Blaha, K. Schwarz, Intern. J. Quantum Chem. **23**, 1535 (1983).
55. P. Blaha, K. Schwarz and P.H. Dederichs, Phys. Rev. B **37**, 2792 (1988).
56. V. Heine, R. Haydock and M. Kelly, in *Solid State Physics, Vol. 35* (H. Ehrenreich, F. Seitz and D. Turnbull, Eds.), Academic, New York (1980).
57. A. Gibson, R. Haydock and J.P. LaFemina, Phys. Rev. B **47**, 9229 (1993).
58. R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).
59. D. Remler and P.A. Madden, Molec. Phys. **70**, 921 (1990).
60. C. Lee, D. Vanderbilt, K. Laasonen, R. Car and M. Parrinello, Phys. Rev. B **47**, 4863 (1993).
61. R. Jastrow, Phys. Rev. **98**, 1479 (1955).
62. S. Fahy, X.W. Wang, G. Louie, Phys. Rev. B **42**, 3503 (1990).
63. X.-P. Li, D.M. Ceperley and R.M. Martin, Phys. Rev. B **44**, 10929 (1991).
64. W. Yang, Phys. Rev. Lett. **66**, 438 (1991); S. Baroni and P. Giannozzi, Europhys. Lett. **17**, 547 (1992); G. Galli and M. Parrinello, Phys. Rev. Lett. **69**, 3547 (1992); E.B. Stechel, A.R. Williams and P.J. Feibelman, Phys. Rev. B **49**, 10088 (1994); W. Yang and T.S. Lee, J. Chem. Phys. **103**, 5674 (1995).
65. G.E.W. Bauer, Phys. Rev. B **27**, 5912 (1983).
66. M. Causà, R. Dovesi and C. Roetti, Phys. Rev. B **43**, 11937 (1991).
67. R. Dovesi, C. Roetti, C. Freyria-Fava, E. Aprà, V.R. Saunders and N.M. Harrison Phil. Trans. R. Soc. London **A341**, 203 (1992).
68. Y.-N. Xu and W.Y. Ching, Phys. Rev. B **48**, 4335 (1993).
69. J. Häglund, G. Grimvall, T. Jarlborg and A. Fernández Guillermet, Phys. Rev. B **43**, 14400 (1991).