

Contents

Introduction	11
1 Theory of electronic structure	13
1.1 The Schrödinger equation	14
1.1.1 Variational formulation	17
1.1.2 The Hellmann-Feynman theorem	19
1.2 Density Functional Theory	21
1.2.1 The Hohenberg-Kohn theorems	22
1.2.2 The Levy approach	25
1.2.3 The representability problem	27
1.2.4 Density operators	29
1.2.5 Canonical-ensemble formulation	32
1.3 The Kohn-Sham mapping	34
1.3.1 Fractional occupancies, Janak's theorem	40
1.3.2 Finite-temperature extension	43
2 Methods of electronic structure	47
2.1 The Local Density Approximation	48
2.2 The pseudopotential approximation	51

2.2.1	Basis representation	52
2.2.2	Smoothness and transferability	53
2.2.3	The Kleinman-Bylander representation	57
2.3	Minimizations and dynamics	59
2.3.1	Car-Parrinello molecular dynamics	60
2.3.2	Conjugate-gradient minimization	64
2.3.3	Metallic systems	71
3	Ensemble Density Functional Theory	83
3.1	Ensemble free energy	84
3.2	Minimization strategy	88
3.2.1	Doubly-preconditioned all-bands CG	89
3.2.2	Occupation matrix: an iterative scheme	96
3.3	Applications and comparisons	100
3.3.1	Self-consistency and molecular dynamics	100
3.3.2	Ionic relaxations	107
3.4	Further developments	109
4	Generalized entropy and cold smearing	113
4.1	Generalized free energy	114
4.2	Smearred functionals	118
4.3	Entropy corrections	121
4.3.1	Energy	123
4.3.2	Forces and other derivatives	126
4.4	Cold smearing	130

5 Ensemble-DFT molecular dynamics	143
5.1 Molecular dynamics simulations	144
5.1.1 Technical details	145
5.2 Metal surfaces at finite temperature	150
5.3 Computational LEED	152
5.3.1 Bulk Al	152
5.3.2 Al(110): thermal relaxations	152
5.3.3 Al(110): layer-resolved displacements	153
5.3.4 Al(111)	155
5.4 Sampling the phase space	156
5.4.1 Preroughening, diffusion and liquid layering	156
Conclusions	171
Acknowledgments	173
Bibliography	174

Chapter 1

Theory of electronic structure

Introduction

A brief review is given of the problem of finding the ground state for a system of interacting electrons immersed in an external potential. The fundamental law of quantum mechanics—the Schrödinger equation [1]—is introduced, and its equivalence to a variational principle is established. A modern and extremely successful reformulation of the problem—that goes under the name of Density Functional Theory [2]—is presented, and its intrinsic nature as a minimization problem is highlighted. The role of temperature [3] and the introduction of statistical operators to describe mixed states is discussed. Finally, the problem is recast via the Kohn-Sham [4] mapping onto a system of non-interacting electrons in a self-consistent external potential.

1.1 The Schrödinger equation

At the present state of human knowledge, quantum mechanics can be regarded as the fundamental theory of atomic phenomena.

[L. I. Schiff, *Quantum Mechanics*, 1955]

This statement has been supported since the early days that followed the introduction of quantum mechanics by an ever growing evidence for the accuracy of its predictions when compared with experiments. Notwithstanding the existence of many unresolved questions in the formal interpretation of the theory [5], the range of its applications has extended from the description of single isolated atoms to all areas of fundamental and applied physics and chemistry. As a simple example of its extreme accuracy, the gyromagnetic ratio for the electron can be determined in agreement with experiment with a precision greater than one part in a billion. Other disciplines, ranging from materials science to biology, are starting now to benefit from this predictive power and from the development of very powerful electronic structure methods, as described in Chapter 2, that allow for the first time to compute fundamental properties of complex systems from first principles, using quantum mechanics to determine the behaviour of their fundamental constituents, electrons and ions.

In the following we are concerned with the ground-state electronic properties of a finite, isolated system of N interacting electrons in an external potential. The external potential considered is that generated by a configuration of atomic nuclei, assumed for the time being to be fixed point charges. Some of these conditions will be relaxed or modified along the way, namely

by introducing pseudopotentials [6][7] to describe the interactions between the valence electrons and the rest of the system, by allowing for the ions to move adiabatically [8] in the field of the electrons' ground state, and by introducing periodic boundary conditions to reduce or eliminate finite-size errors.

The non-relativistic time-independent Schrödinger equation for the system described is

$$\hat{H}\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N) = E\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N), \quad (1.1)$$

where the Hamiltonian operator¹ \hat{H} is

$$\begin{aligned} \hat{H} &= \hat{T}_e + \hat{V}_{n-e} + \hat{U}_{e-e} + \hat{W}_{n-n} = \\ &= \sum_i \left(-\frac{1}{2}\nabla_i^2 + v(\mathbf{r}_i) \right) + \frac{1}{2} \sum_{i \neq j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \hat{W}_{n-n} \end{aligned} \quad (1.2)$$

with

$$v(\mathbf{r}) = - \sum_{\alpha} \frac{Z_{\alpha}}{|\mathbf{r} - \mathbf{R}_{\alpha}|} \quad \text{and} \quad W_{n-n} = \frac{1}{2} \sum_{\alpha\beta} \frac{Z_{\alpha}Z_{\beta}}{R_{\alpha\beta}}. \quad (1.3)$$

The terms in (1.2) are associated, respectively, with the kinetic energy T_e of the electrons, the potential energy V_{n-e} of the electrons in the field of α nuclei of charge Z_{α} , the electrostatic energy U_{e-e} between the electrons, and the electrostatic energy W_{n-n} between the nuclei.

Equation (1.1) is an eigenvalue equation for the N -electron many-body wavefunction $\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N)$ ($\mathbf{r}_1 \dots \mathbf{r}_N$ being the spatial coordinates for the electrons), where \hat{H} is Hermitian; in order for the solutions to be acceptable they must belong to the Hilbert space $\mathcal{L}^2(\mathbb{R}^3) \otimes \dots \otimes \mathcal{L}^2(\mathbb{R}^3)$ of square-integrable functions. The spin coordinates will be ignored (but not the spin=1/2 nature of the electrons), while some relativistic effects—relevant

¹Atomic units are used: $\hbar = m_e = e = 4\pi\epsilon_0 = 1$.

for heavy atoms—can be taken into account a posteriori when constructing the pseudopotential. The Fermi-Dirac statistics for the electrons enters the problem as a restriction on the Hilbert space, requiring Ψ to be antisymmetric under the exchange of the coordinates of two electrons. This, translating into the Pauli principle, accounts ultimately for the stability of matter, with the ground state energy bounded below by a constant times the first power of the particle number [9][10].

The solutions to equation (1.1) form a complete set $\{\Psi_n\}$ [11]; once normalized, any two of them that correspond to different eigenvalues are orthogonal, and the set of eigenvectors $\{\Psi_n\}$ is thus always considered fully orthonormal. According to the postulates of quantum mechanics [12], to each observable A there corresponds a Hermitian operator \hat{A} and consequently its complete set of orthonormal eigenfunctions $\{\phi_n\}$ with eigenvalues $\{a_n\}$; the wavefunction of the system can then always be expressed in terms of these eigenfunctions as

$$\Psi = \sum_n c_n \phi_n. \quad (1.4)$$

If we have an ensemble of systems identically prepared, each measurement will yield the eigenvalue a_n with probability $|c_n|^2$. The expectation value of A can thus be expressed, using Dirac notation, as

$$\langle \hat{A} \rangle = \langle \Psi | \hat{A} | \Psi \rangle = \sum_n \sum_{n'} c_{n'}^* c_n \langle \phi_{n'} | \hat{A} | \phi_n \rangle = \sum_n |c_n|^2 a_n, \quad (1.5)$$

where $\langle \Psi | \hat{A} | \Psi \rangle$ stands for $\int \Psi^* \hat{A} \Psi \, d\mathbf{r}_1 \dots d\mathbf{r}_N$.

1.1.1 Variational formulation

Let us consider a time-independent Hamiltonian, as that described in (1.2); for simplicity, we assume that the system is either enclosed in a box of finite size, or that periodic boundary conditions² are applied. Under such conditions, the (non-empty) spectrum of eigenvalues $\{E_n\}$ and eigenfunctions $\{\psi_n\}$ is discrete³ (see Chap. 3 of Ref. [13] for a discussion). For an arbitrary function Ψ in the Hilbert space $\mathcal{L}^2(\mathbb{R}^3) \otimes \dots \otimes \mathcal{L}^2(\mathbb{R}^3)$ that has non-vanishing norm, we define the functional $E[\Psi]$ as

$$E[\Psi] = \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \quad (1.6)$$

The following theorem is easily proved:

Variational Principle *The Schrödinger equation $\hat{H}\Psi = E\Psi$ for the Hamiltonian \hat{H} is equivalent to the variational principle $\delta E[\Psi] = 0$.*

Proof: Let us consider a generic unconstrained Ψ ; taking the variation $\delta(E[\Psi] \langle \Psi | \Psi \rangle)$ in (1.6) we have:

$$\delta E[\Psi] \langle \Psi | \Psi \rangle + E[\Psi] \langle \delta \Psi | \Psi \rangle + E[\Psi] \langle \Psi | \delta \Psi \rangle = \langle \delta \Psi | \hat{H} | \Psi \rangle + \langle \Psi | \hat{H} | \delta \Psi \rangle.$$

²In electronic structure calculations it is useful to adopt a looser form of periodic boundary conditions that goes under the names of Born and von Karman: $\Psi(\mathbf{r} + N_i \mathbf{a}_i) = \Psi(\mathbf{r})$, $i = 1, 2, 3$, where \mathbf{a}_i are the primitive vectors of the cell and N_i are the independent primitive cells in each direction. These conditions impose the existence of a discrete, finite number of allowed values for the Bloch wave vector (a good quantum number in a periodic solid) and its corresponding energy. The spacing between different wavevectors goes to zero in the limit $N_i \rightarrow \infty$, recovering thus the continuum limit for the spectrum of a truly infinite periodic crystal.

³The discussion can be generalized to an Hamiltonian with an additional continuum spectrum $\{e_n\}$ such that $\inf\{e_n\} > \min\{E_n\}$.

The following equivalences are established, exploiting the hermiticity of \hat{H} :

$$\delta E[\Psi] = 0 \Leftrightarrow \langle \delta\Psi | \hat{H} - E | \Psi \rangle + \langle \Psi | \hat{H} - E | \delta\Psi \rangle = 0 \Leftrightarrow (\hat{H} - E[\Psi]) | \Psi \rangle = 0$$

QED

The variational principle can be reformulated in order to include automatically the wavefunction normalization, by introducing an undetermined Lagrange multiplier E : the constrained variational problem

$$\delta[\langle \Psi | \hat{H} | \Psi \rangle] = 0 \quad , \quad \langle \Psi | \Psi \rangle = 1$$

is thus transformed in the unconstrained variational problem

$$\delta[\langle \Psi | \hat{H} | \Psi \rangle - E \langle \Psi | \Psi \rangle] = 0. \quad (1.7)$$

The importance of the functional (1.6) is that it provides a variational bound to the ground-state energy of the system; this can be seen by expanding Ψ in the energy eigenvectors $\{\psi_n\}$:

$$\Psi = \sum_n a_n \psi_n \Rightarrow E[\Psi] = \frac{\sum_n |a_n|^2 E_n}{\sum_n |a_n|^2} \Rightarrow E[\Psi] - E_0 = \frac{\sum_n |a_n|^2 (E_n - E_0)}{\sum_n |a_n|^2}$$

where we have indicated with $E_0 = \min\{E_n\}$ the ground-state energy. It is clear from this that a trial Ψ provides an upper bound to the real ground-state energy, and that the value E_0 is reached if and only if $\Psi = \alpha \Psi_0, \alpha \in \mathbb{C}$. The fundamental *minimum principle* thus follows:

$$E[\Psi] \geq E_0 \quad ; \quad \Psi = \alpha \Psi_0 \iff E[\Psi] = E_0 \quad (1.8)$$

The ground-state energy E implicitly defined by the minimization in (1.8) depends only on and is completely identified by the choice of N and v in (1.2);

in this sense, E is said to be a *functional* $E[N, v]$ of the particle number N and of the external potentials v .

This minimum principle has a key importance in practical applications: as an example, it is the basis of the widely used *Rayleigh-Ritz variational method*, where a trial function is constructed that depends on a number of variational parameters, that are determined by the requirement of minimizing the expectation value for the energy. Obviously, the quality of the results depends both on the choice of the trial function and on the dimensionality of the parameter space that is spanned. In addition, if a trial function can be constructed as to be orthogonal to the $n - 1$ lowest eigenfunctions, the Rayleigh-Ritz method provides an upper bound for the n -th eigenvalue. The well-known *Hartree* [14] and *Hartree-Fock* [15] methods correspond to searching the solutions for (1.7) respectively in the subspace of the products of single-particle orbitals and in the subspace of antisymmetrized products of single-particle orbitals (the Slater determinants).

1.1.2 The Hellmann-Feynman theorem

There is an important consequence that can be drawn from equation (1.7): if the Hamiltonian H depends parametrically on some λ , such that $d\hat{H}_\lambda/d\lambda$ is well defined, the following theorem can be demonstrated:

Hellmann-Feynman theorem *Given a Hamiltonian \hat{H}_λ , and a $\Psi(\lambda)$ satisfying the Schrödinger equation $\hat{H}_\lambda\Psi = E_\lambda\Psi$, the following relation holds:*

$$\frac{dE_\lambda}{d\lambda} = \frac{d}{d\lambda} \langle \Psi | \hat{H}_\lambda | \Psi \rangle = \langle \Psi | d\hat{H}_\lambda / d\lambda | \Psi \rangle$$

Proof: It follows from the orthonormality property for the solutions of the

Schrödinger equation, or its equivalent variational formulation (1.7), that:

$$\langle \Psi(\lambda) | \Psi(\lambda) \rangle = 1 \Rightarrow \frac{d}{d\lambda} \langle \Psi | \Psi \rangle = 0 \Rightarrow \langle d\Psi/d\lambda | \Psi \rangle + \langle \Psi | d\Psi/d\lambda \rangle = 0.$$

Using the hermiticity of \hat{H} , it is readily shown that:

$$\begin{aligned} \frac{d}{d\lambda} \langle \Psi | \hat{H}_\lambda | \Psi \rangle &= \langle d\Psi/d\lambda | \hat{H}_\lambda | \Psi \rangle + \langle \Psi | d\hat{H}_\lambda/d\lambda | \Psi \rangle + \langle \Psi | \hat{H}_\lambda | d\Psi/d\lambda \rangle = \\ &= E(\langle d\Psi/d\lambda | \Psi \rangle + \langle \Psi | d\Psi/d\lambda \rangle) + \langle \Psi | d\hat{H}_\lambda/d\lambda | \Psi \rangle = \langle \Psi | d\hat{H}_\lambda/d\lambda | \Psi \rangle \end{aligned}$$

QED

The scope of the theorem is trivially extended to the expectation values of any Hermitian operator \hat{A}_λ dependent on some external parameter λ and for a system prepared in a given eigenstate of \hat{A}_λ .

The original work [16] was concerned with the definition and the calculation of the forces acting on molecular systems: in fact Hamiltonians \hat{H} such as that defined in (1.2) explicitly depend on the nuclear coordinates \mathbf{R}_α , and it is much more convenient to calculate directly the expectation value of $d\hat{H}/d\mathbf{R}_\alpha$ than approximating this derivative with the energy differences⁴. Feynman's original application of the theorem to the Hamiltonian (1.2) leads immediately to the *electrostatic theorem*: the force on a nucleus is given by a completely classical expression, where the potential is obtained by the superposition of the nuclear fields and by the electrostatical potential generated by the electronic charge density $n(\mathbf{r}) = N \int |\Psi(\mathbf{r}_1 \dots \mathbf{r}_N)|^2 d\mathbf{r}_2 \dots d\mathbf{r}_N$:

$$\begin{aligned} \mathbf{F}_\alpha &= -\nabla_\alpha W_{n-n} - \sum_i \int |\Psi(\mathbf{r}_1 \dots \mathbf{r}_N)|^2 \nabla_\alpha v(\mathbf{r}_i) d\mathbf{r}_1 \dots d\mathbf{r}_N = \\ &= -\nabla_\alpha W_{n-n} - \int n(\mathbf{r}) \nabla_\alpha v(\mathbf{r}) d\mathbf{r}. \end{aligned} \quad (1.9)$$

⁴A similar problem for a time-dependent Schrödinger equation had already been addressed by Ehrenfest [17] in 1927; namely he demonstrated a quantum mechanical equivalent $m d^2 \langle \Psi | \mathbf{r} | \Psi \rangle / dt^2 = \langle \Psi | -dV/d\mathbf{r} | \Psi \rangle$ for Newton's second law of dynamics.

1.2 Density Functional Theory

The solution to the electronic structure problem has been presented— according to the prescriptions of the Schrödinger equation— as the problem of obtaining the proper ground-state many-body wavefunction $\Psi_v(\mathbf{r}_1, \dots, \mathbf{r}_N)$ for a given external potential v . As such, the ground-state solution has an intrinsically high complexity: while the problem is completely defined by the total numbers of particles N and by the external potential $v(\mathbf{r})$, its solution— uniquely determined by N and $v(\mathbf{r})$, and thus a functional $\Psi \equiv \Psi[N, v(\mathbf{r})]$ — depends on $3N$ coordinates. This makes the direct search for either exact or approximate solutions to the many-body problem a task of rapidly increasing complexity.

Soon after the introduction of the Schrödinger equation there have been attempts to refocus the search on the ground-state charge density $n(\mathbf{r})$, which is the most natural and physical low-complexity quantity in the problem. The first outcome of this effort was the Thomas-Fermi model [18][19]: a functional of the charge density alone was constructed from the basic assumptions of treating the electrons as independent particles, reducing the electron-electron interaction to the Coulomb electrostatic energy, and using a *local density approximation* for the kinetic energy term:

$$T[n] = \int t[n(\mathbf{r})]d\mathbf{r},$$

where $t[n(\mathbf{r})]$ is the kinetic energy density for a system of non-interacting electrons with density n . The functional that is obtained from this hypothesis

is⁵

$$E_{TF}[n(\mathbf{r})] = \frac{3}{10}(3\pi^2)^{\frac{2}{3}} \int n^{\frac{5}{3}}(\mathbf{r})d\mathbf{r} - \int v(\mathbf{r})n(\mathbf{r}) + \frac{1}{2} \int \int \frac{n(\mathbf{r}_1)n(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2$$

and the Thomas-Fermi solution is defined as the charge density $n(\mathbf{r})$ that minimizes this functional E_{TF} , with the constraints of being greater than zero and normalized to the total number of electrons N . This approach has been shown to fail in providing even qualitatively correct descriptions of systems more complex than isolated atoms [20], its main limitation coming from the approximate form for the kinetic energy contribution (it should be noted that more sophisticated and accurate choices for the kinetic energy term as a function of the electron density can be made [21]). On a more fundamental basis, the model was introduced when there was no theoretical ground to support the choice of the charge density, rather than the many-body wavefunction, for the search, via the variational principle (1.8), of the electronic ground state.

1.2.1 The Hohenberg-Kohn theorems

The essential role that is played by the charge density in the search for the electronic ground state was pointed out for the first time by Hohenberg and Kohn [2] in their exact reformulation of the problem that now goes under the name of *Density Functional Theory*.

Let us consider a finite, isolated system of N interacting electrons in an external potential $v(\mathbf{r})$, as described in Sec. 1.1, with a Hamiltonian (1.2); the only additional assumption is that the ground-state Ψ be unique, non-degenerate. This latter assumption will be relaxed in 1.2.2. Again, Ψ and

⁵Neglecting for the rest of the chapter contributions from the term W_{n-n} .

$n(\mathbf{r}) = \langle \Psi | \hat{n} | \Psi \rangle$ are perfectly specified (albeit not explicitly) once N and $v(\mathbf{r})$ are formulated, and thus are *functionals* of $[N, v(\mathbf{r})]$. Let now $\mathcal{V}_N(\mathbb{R}^3)$ be the class of all densities that are positive definite, normalized to an integer number N , and such that *there exists* an external potential $v(\mathbf{r})$ for which there is a non-degenerate ground state corresponding to that density; this is the class of v -representable densities, and the following discussion⁶ will be restricted to densities that belong to this class. The first theorem of Hohenberg and Kohn [2] establishes the legitimacy of the charge density as the fundamental variable in the electronic problem.

1st Theorem: the Density as the Basic Variable $\forall n(\mathbf{r}) \in \mathcal{V}_N(\mathbb{R}^3)$ *the external potential $v(\mathbf{r})$ is a functional $v(\mathbf{r}) \equiv v(\mathbf{r})[n(\mathbf{r})]$ of the charge density $n(\mathbf{r})$, within an additive constant.*

Proof: Let us assume, *ad absurdum*, that there exists a different potential v' with a ground state Ψ' that gives rise to the same charge density $n(\mathbf{r})$. Let E and E' be the respective ground-state energies; taking Ψ' as a trial solution for the Hamiltonian \hat{H} , we have the strict inequality

$$E < \langle \Psi' | \hat{H} | \Psi' \rangle = \langle \Psi' | \hat{H}' | \Psi' \rangle + \langle \Psi' | \hat{H} - \hat{H}' | \Psi' \rangle = E' + \int n(\mathbf{r})[v(\mathbf{r}) - v'(\mathbf{r})] d\mathbf{r}$$

(if $\langle \Psi' | \hat{H} | \Psi' \rangle = E$ then Ψ' would be the ground state for H , in virtue of (1.8), but the two distinct differential equations for \hat{H} and \hat{H}' cannot have the same ground state). Similarly, taking Ψ as a trial solution for \hat{H}' , we have

$$E' < \langle \Psi | \hat{H}' | \Psi \rangle = \langle \Psi | \hat{H} | \Psi \rangle + \langle \Psi | \hat{H}' - \hat{H} | \Psi \rangle = E - \int n(\mathbf{r})[v(\mathbf{r}) - v'(\mathbf{r})] d\mathbf{r}$$

⁶See 1.2.3 for the problem of v - and N -representability of charge densities and density matrices.

Adding together these two inequalities the *absurdum* is obtained:

$$E + E' < E + E'.$$

QED

It is thus demonstrated that the ground-state charge density determines, in principle, the external potential of the Schrödinger equation of which it is solution⁷ and thus, implicitly, all the other electronic ground-state properties of the system. It is then possible to introduce, in the class of v -representable densities $\mathcal{V}_N(\mathbb{R}^3)$, the functional

$$F[n(\mathbf{r})] = \langle \Psi | \hat{T}_e + \hat{U}_{e-e} | \Psi \rangle; \quad (1.10)$$

this is unique and well defined, and, since it does not explicitly depend on the external potential for the electrons, it is a *universal* functional of the charge density. It should be pointed out that no explicit expressions for this functional are known to date, obviously as a by-product of the complexity of the many-body problem that lies at the core of the definition of $F[n(\mathbf{r})]$ [22].

It is now possible to define, for an *arbitrary* external potential $v(\mathbf{r})$ (assumed here to be local), unrelated to the external potential implicitly defined by a density in $\mathcal{V}_N(\mathbb{R}^3)$, the Hohenberg-Kohn *Density Functional* for the energy of the ground state:

$$E_v[n'(\mathbf{r})] = F[n'(\mathbf{r})] + \int v(\mathbf{r})n'(\mathbf{r})d\mathbf{r}. \quad (1.11)$$

The second theorem of Hohenberg and Kohn [2] establishes the existence of a variational principle for this functional of the charge density, thus rationalizing the original intuition of Thomas and Fermi.

⁷Precisely, it is an expectation value of the solution, since $n(\mathbf{r})$ determines $v(\mathbf{r})$, which in turn determines the ground state Ψ , and the relation $n(\mathbf{r}) = \langle \Psi | \hat{n} | \Psi \rangle$ holds.

2nd Theorem: the Energy Variational Principle $\forall n'(\mathbf{r}) \in \mathcal{V}_N(\mathbb{R}^3)$, $E_v[n'(\mathbf{r})] = F[n'(\mathbf{r})] + \int v(\mathbf{r})n'(\mathbf{r})d\mathbf{r} \geq E_0$, where E_0 is the ground-state energy for N electrons in the external potential $v(\mathbf{r})$.

Proof: A given $n'(\mathbf{r})$ determines uniquely, via the 1st theorem, its own external potential and ground-state wavefunction Ψ' . If *this* Ψ' is used as a trial wavefunction for the expectation value of the Hamiltonian with the external potential $v(\mathbf{r})$, the relation $\langle \Psi' | \hat{H} | \Psi' \rangle \equiv \langle \Psi' | \hat{T}_e + \hat{U}_{e-e} | \Psi' \rangle + \langle \Psi' | \hat{v} | \Psi' \rangle = F[n'(\mathbf{r})] + \int v(\mathbf{r})n'(\mathbf{r})d\mathbf{r} = E_v[n'(\mathbf{r})]$ holds. Now, because of the minimum principle 1.8, $E_v[n'(\mathbf{r})] = \langle \Psi' | \hat{H} | \Psi' \rangle \geq E_0$. Since the ground state is non-degenerate, the equality holds only if Ψ' is the ground state for the potential v .

QED

These two theorems show that the problem of solving the Schrödinger equation for the ground state can be exactly recast into the variational problem of minimizing the Hohenberg-Kohn functional (1.11) with respect to the charge density; the complexity of the problem is reduced in principle from having to deal with a function of $3N$ variables to one that depends only on the 3 spatial coordinates.

1.2.2 The Levy approach

A simpler and more direct proof for the Hohenberg-Kohn theorems has been introduced by Levy [23]; this has the advantage of removing the requirement of non-degeneracy for the ground state of the electronic problem, while focusing directly on the minimization procedure in the definition of the func-

tional itself. In addition, the ill-defined problem of the representability of the charge density is greatly simplified (see 1.2.3).

Levy [23][24] introduced a new definition for functionals of the density, based on the concept of *constrained search*. Given an operator \hat{A} and a density $n(\mathbf{r})$, the functional $A[n(\mathbf{r})]$ is defined as the minimum expectation value of A for a constrained search along all the many-body antisymmetric wavefunctions for which $\langle \Psi | \hat{n} | \Psi \rangle = n(\mathbf{r})$:

$$A[n(\mathbf{r})] = \min_{\Psi \rightarrow n} \langle \Psi | \hat{A} | \Psi \rangle \quad (1.12)$$

This definition is meaningful (the set is non-empty) for those densities that do admit a representation derived from a N -body wavefunction; it is the class of N -representable densities, much simpler to characterize than the class of v -representable densities (see 1.2.3). The set is also known to be bounded from below [25]. The Levy proof develops as follows⁸:

Theorem: the Levy proof *The ground-state electron density $n_0(\mathbf{r})$ minimizes the functional $E_v[n(\mathbf{r})] = F[n(\mathbf{r})] + \int v(\mathbf{r})n(\mathbf{r})d\mathbf{r}$, where $F[n(\mathbf{r})]$ is a universal functional of $n(\mathbf{r})$; the minimum value is the ground-state electronic energy E_0 .*

Proof: Let the operator \hat{F} be $\hat{T}_e + \hat{U}_{e-e}$; $F[n(\mathbf{r})]$ is defined uniquely (and it is a universal functional of the charge density alone) via the constrained search $\min_{\Psi \rightarrow n} \langle \Psi | \hat{F} | \Psi \rangle$. For an arbitrary $n(\mathbf{r})$, let Ψ_n be the N -electron wavefunction that yields the charge density $n(\mathbf{r})$ and minimizes $\langle \Psi | \hat{F} | \Psi \rangle$. $E_v[n(\mathbf{r})]$ is $F[n(\mathbf{r})] + \int v(\mathbf{r})n(\mathbf{r})d\mathbf{r}$, and so it follows from the definition of Ψ_n that $E_v[n(\mathbf{r})] = \langle \Psi_n | \hat{F} + \hat{v} | \Psi_n \rangle$; but this is the Hamiltonian in the presence

⁸As a matter of clarity, a subscript is introduced to denote all the quantities that refer to the ground state (Ψ_0, n_0, E_0).

of the external potential $v(\mathbf{r})$, and so we have, from the minimum principle (1.8), that for every N -representable charge density

$$E_v[n(\mathbf{r})] \geq E_0.$$

Now it remains to be demonstrated what is the value of E_v evaluated on the ground state charge density n_0 . Since the ground-state wavefunction Ψ_0 integrates to the ground-state charge density n_0 , it certainly takes part in the constrained search that defines $F[n_0(\mathbf{r})]$. Thus, $F[n_0(\mathbf{r})] \leq \langle \Psi_0 | \hat{F} | \Psi_0 \rangle$, and so, adding the external potential term on both sides, we have

$$E_v[n_0(\mathbf{r})] \leq E_0,$$

that, together with the previous relation, gives us $E_v[n_0(\mathbf{r})] = E_0$.

QED

It should be mentioned that, although very powerful, this approach misses one of the elegant results of Hohenberg and Kohn, namely that there are no two different external potentials (a part from some additive constant) that give rise to the same ground-state electronic charge density.

1.2.3 The representability problem

The original Hohenberg-Kohn proof takes place in the space of charge densities that are v -representable; again, this means that they are positive definite, normalized to an integer number, and such that there exists an external potential $v(\mathbf{r})$ for which there is a non-degenerate ground state corresponding to that density. In this framework, the Hohenberg-Kohn 1st theorem establishes a one-to-one mapping between a v -representable charge density and

the external potential for which this density is the the expectation value of the ground state, and in this space $\mathcal{V}_N(\mathbb{R}^3)$ the universal functional F and the variational principle of the 2^{nd} theorem are formulated. The characterization of the class $\mathcal{V}_N(\mathbb{R}^3)$ is still lacking; v -representability has been demonstrated, in a lattice version of the Schrödinger equation, for all the particle-conserving charge densities that are in the neighbour of a v -representable density [26]. This, together with the more serious necessity to adopt approximations for the universal functional F , has lead to a generalized application of the variational principle for all of its practical implementations. It should be pointed out that simple counter-examples of densities that do not belong to $\mathcal{V}_N(\mathbb{R}^3)$ have been presented [24][27]; e.g. if a Hamiltonian has a ground state with degeneracy greater than 2, the density obtained from a linear average of each ground state is in general not v -representable [24].

The Levy proof overcomes this problem, by requiring the density just to be obtainable from some antisymmetric wavefunction (in addition to being positive definite and normalized to an integer number), in order to allow for a meaningful constrained search. This is the condition of N -representability, a much weaker requirement that is satisfied by any well-behaved charge density. In fact, it has been demonstrated by Gilbert [28] that the only condition required is proper differentiability,

$$\int |\nabla n^{\frac{1}{2}}(\mathbf{r})|^2 d\mathbf{r} \in \mathbb{R},$$

to ensure that the kinetic energy of the auxiliary orbitals used in the construction of an antisymmetric wavefunction from a given $n(\mathbf{r})$ remains finite.

1.2.4 Density operators

Up to now, a quantum state has been described in terms of its N-body wavefunction $\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N)$, with the expectation value for a given observable \hat{A} (Hermitian operator) given by $\langle \hat{A} \rangle = \langle \Psi | \hat{A} | \Psi \rangle$. A more general description can be obtained with the introduction of the *density operator* $\hat{\gamma}_N = |\Psi\rangle\langle\Psi|$, whose representation in the coordinate space is the *density matrix*

$$\Psi^*(\mathbf{r}'_1, \dots, \mathbf{r}'_N) \Psi(\mathbf{r}_1, \dots, \mathbf{r}_N).$$

For a normalized Ψ , $\text{tr}(\hat{\gamma}_N) \equiv \int \Psi^*(\mathbf{r}_1, \dots, \mathbf{r}_N) \Psi(\mathbf{r}_1, \dots, \mathbf{r}_N) d\mathbf{r}_1 \dots d\mathbf{r}_N = 1$; $\hat{\gamma}_N$ is clearly idempotent, and thus it is a projection operator. The expectation values can now be expressed as a trace on this operator; in the coordinate representation and for a normalized Ψ we have

$$\langle \hat{A} \rangle = \text{tr}(\hat{\gamma}_N \hat{A}) = \text{tr}(\hat{A} \hat{\gamma}_N) = \langle \Psi | \hat{A} | \Psi \rangle. \quad (1.13)$$

The description of a system at a finite temperature requires the introduction of an *ensemble density operator*, since then the system is part of a much larger closed system (a heat bath with quasi-infinite thermal capacity) for which a complete Hamiltonian description is unattainable. In this case, the system is said to be in a *mixed state*, where it cannot be characterized by a wavefunction (as opposed to the *pure states* of all the preceding discussion), but only as a mixture over all pure states. The ensemble density operator $\hat{\Gamma}_N$ is thus defined as $\hat{\Gamma}_N = \sum_i w_i |\Psi_i\rangle\langle\Psi_i|$, where the sum is taken over all the available pure states. The w_i are often described as the probabilities of finding the system in the relative pure state, although this is conceptually misleading [29], and the w_i are just a statistical distribution outcome of both the probabilistic interpretation of the wavefunction and of the statistical description

coming for our imperfect knowledge of the full system. This becomes apparent in the statistical generalization of (1.13) for the expectation value of an observable for a mixed state, which is given by

$$\langle \hat{A} \rangle = \text{tr}(\hat{\Gamma}_N \hat{A}) = \text{tr}(\hat{A} \hat{\Gamma}_N) = \sum_i w_i \langle \Psi_i | \hat{A} | \Psi_i \rangle, \quad (1.14)$$

where the cross terms that would arise if the mixed state was actually a superposition of pure states are not included. In contrast to the density operator, the ensemble density operator is no longer idempotent; they are both Hermitian and uniquely defined (as opposed to the wavefunction, that has an arbitrary phase attached). An important relation follows from the definitions and from the time-dependent Schrödinger equation $i\hbar|\dot{\Psi}\rangle = \hat{H}|\Psi\rangle$:

$$\frac{\partial}{\partial t} \hat{\gamma}_N = \left(\frac{\partial}{\partial t} |\Psi\rangle \right) \langle \Psi| + |\Psi\rangle \left(\frac{\partial}{\partial t} \langle \Psi| \right) = \frac{\hat{H}}{i\hbar} |\Psi\rangle \langle \Psi| - |\Psi\rangle \langle \Psi| \frac{\hat{H}}{i\hbar},$$

from which the *Liouville equations* for the density operator and, by the linearity of $\hat{\Gamma}_N = \sum_i w_i |\Psi_i\rangle \langle \Psi_i|$, for the ensemble density operator, are obtained:

$$i\hbar \frac{\partial}{\partial t} \hat{\gamma}_N = [\hat{H}, \hat{\gamma}_N], \quad i\hbar \frac{\partial}{\partial t} \hat{\Gamma}_N = [\hat{H}, \hat{\Gamma}_N]. \quad (1.15)$$

The Hamiltonian and the ensemble density operator commute in the case of a stationary state:

$$[\hat{H}, \hat{\Gamma}_N] = 0, \quad (1.16)$$

and a complete system of eigenvectors for both of them can be found. This is a fundamental relation that will be fully exploited in the development of a practical algorithm (albeit for the slightly different case of non-interacting electrons in a self-consistent potential).

It is possible and useful to define a *reduced density matrix of p^{th} order*, which in the coordinate representation has the following form:

$$\gamma_p(\mathbf{r}'_1 \dots \mathbf{r}'_p; \mathbf{r}_1 \dots \mathbf{r}_p) = \frac{N!}{p! (N-p)!} \int \dots \int \Psi^*(\mathbf{r}'_1 \dots \mathbf{r}'_N) \Psi(\mathbf{r}_1 \dots \mathbf{r}_N) d\mathbf{r}_{p+1} \dots d\mathbf{r}_N.$$

Since the Hamiltonian \hat{H} in (1.2) is a function of one-particle and two-particle terms only, it can be demonstrated that its expectation value can be written as an *explicit functional* of the density matrices of order 1 and 2 [30] (Sec. 2.3), or, since γ_1 descends from γ_2 , just of the density matrix of order 2:

$$\begin{aligned} E &= \text{tr}(\hat{H}\hat{\gamma}_N) = E[\gamma_1, \gamma_2] = E[\gamma_2] = \\ &= \int [(-\frac{1}{2}\nabla^2 + v(\mathbf{r})) \gamma_1(\mathbf{r}, \mathbf{r}) d\mathbf{r} + \iint \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \gamma_2(\mathbf{r}_1, \mathbf{r}_2; \mathbf{r}_1, \mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2. \end{aligned} \quad (1.17)$$

It is possible to construct a density-matrix functional theory and demonstrate a variational principle for the functional in (1.17) [31], with the eigenvectors of γ_1 (the natural orbitals) and γ_2 (the natural geminals) as the search variables, and reducing again the complexity of the problem from a $3N$ -dimensional search to a 6-dimensional search. The great obstacle is that γ_1 and γ_2 must be N -representable, thus descending from an appropriate antisymmetric Ψ , and this characterization (especially for γ_2) is still ⁹ unsolved [32]). It is nevertheless interesting to contrast this problem of minimizing an explicit functional in a class of N -representable density matrices that escapes characterization, with the density-functional problem of minimizing a functional that is defined only implicitly in the well-defined simple class of N -representable densities.

⁹Note that γ_2 depends on 4 independent variables, but, because of density-functional theory, is completely characterized by $n(\mathbf{r})$. This means that the set of N -representable density matrices of order 2 has zero Lebesgue measure.

1.2.5 Canonical-ensemble formulation

From the fundamentals of classical statistical mechanics we have that the entropy is

$$S(E, V) \equiv k_B \ln \Omega(E), \quad (1.18)$$

where $\Omega(E)$ is the volume in the phase space occupied by the microcanonical ensemble at energy E . The same definition holds for quantum mechanics, with $\Omega(E)$ computed properly via a sum over all the states of the system:

$$\Omega(E) = \sum_{\{w_i\}} W[\{w_i\}], \quad (1.19)$$

where $W[\{w_i\}]$ is the number of states available for a given set of occupations w_i of the accessible quantum states, and the sum is done on all the possible sets that are compatible with the thermodynamical ensemble. In the statistical description via the ensemble density operator $\hat{\Gamma}_N = \sum_i w_i |\Psi_i\rangle\langle\Psi_i|$, the distribution is over a set of pure states $|\Psi_i\rangle$, with the w_i following the rules of a distribution ($w_i \geq 0$ and $\sum_i w_i = 1$). The entropy can thus be equivalently written as

$$S(E, V) = -k_B \sum_i w_i \ln w_i = -k_B \text{tr} (\hat{\Gamma}_N \ln \hat{\Gamma}_N), \quad (1.20)$$

where the second equality exploits the independence of the value of the trace of an operator from the representation in which it is expressed.

The choice of the ensemble (canonical, in this case) determines the equilibrium ensemble density operator, by requiring that it maximizes the entropy while keeping the constraints on the statistical distribution $\sum_i w_i = 1$ and on the expectation value for the energy $\text{tr}(\hat{\Gamma}_N \hat{H}) = E$. These conditions lead

to the equilibrium canonical ensemble density operator $\hat{\Gamma}_N^0$ [30] (Sec. 2.7):

$$\hat{\Gamma}_N^0 = \frac{e^{-\beta\hat{H}}}{\text{tr} \left(e^{-\beta\hat{H}} \right)}.$$

Following (1.20) the *Helmholtz free energy* $A = E - TS$ for a given ensemble density operator can be introduced:

$$A[\Gamma_N] \equiv \text{tr} \hat{\Gamma}_N \left(\frac{1}{\beta} \ln \hat{\Gamma}_N + \hat{H} \right). \quad (1.21)$$

Now a variational principle can be demonstrated, that is the exact equivalent of the Rayleigh-Ritz variational principle (1.8) for the case of finite temperature, namely that for every positive definite $\hat{\Gamma}_N$ with unit trace:

$$A[\Gamma_N] \geq A[\Gamma_N^0]. \quad (1.22)$$

where Γ_N^0 is the ground-state ensemble density operator. The proof relies on Gibbs' inequality ($\sum_i w_i (\ln w_i - \ln w_i^0) \geq 0$) and on Jensen's inequality ($\langle \exp(\hat{A}) \rangle \geq \exp(\langle \hat{A} \rangle)$) and can be found in Ref. [33] (Sec. 2.11).

Once the variational principle is formulated, the demonstration of the existence of a finite-temperature, canonical density functional theory follows the steps of Sec. 1.2, either along the original Hohenberg-Kohn proof [3], or in a constrained search formulation. In brief, following the Levy approach, a (canonical ensemble) universal functional of the charge density $n(\mathbf{r})$ can be uniquely defined, via the constrained search on all the density operators that integrate to $n(\mathbf{r})$:

$$F_\beta[n(\mathbf{r})] = \min_{\hat{\Gamma}_N \rightarrow n} \text{tr} \left[\hat{\Gamma}_N \left(\hat{T}_e + \hat{U}_{e-e} + \frac{1}{\beta} \ln \hat{\Gamma}_N \right) \right] \quad (1.23)$$

With this universal functional at hand, the canonical Mermin density functional is written as:

$$A_v[n(\mathbf{r})] = F_\beta[n(\mathbf{r})] + \int v(\mathbf{r})n(\mathbf{r}) \, d\mathbf{r}.$$

Now, for a given $n(\mathbf{r})$, a Γ_N that minimizes (1.23) is defined, and a variational principle for the Mermin functional holds (from Eq. 1.22):

$$A_v[n(\mathbf{r})] = A[\Gamma_N] \geq A[\Gamma_N^0]. \quad (1.24)$$

The Mermin functional evaluated on the ground-state charge density $n_0(\mathbf{r})$ involves a search on all the ensemble density operators that integrate to $n_0(\mathbf{r})$, and thus includes the ground state Γ_N^0 ; it follows that $A_v[n_0(\mathbf{r})] \leq A[\Gamma_N^0]$, that, combined with the variational principle for the Mermin functional, gives $A_v[n_0(\mathbf{r})] = A[\Gamma_N^0]$.

1.3 The Kohn-Sham mapping

Density Functional Theory provides the theoretical ground for reformulating the ground-state many-body problem as a variational problem on the charge density; the constrained minimization on the Hohenberg-Kohn density functional for N electrons can be rewritten, introducing the indeterminate multiplier μ , as the variational problem:

$$\delta \left[F[n(\mathbf{r})] + \int v(\mathbf{r})n(\mathbf{r}) d\mathbf{r} - \mu \left(\int n(\mathbf{r}) d\mathbf{r} - N \right) \right] = 0. \quad (1.25)$$

Formally (1.25) leads to the Euler-Lagrange equation for the charge density¹⁰

$$\frac{\delta F[n(\mathbf{r})]}{\delta n(\mathbf{r})} + v(\mathbf{r}) = \mu. \quad (1.26)$$

¹⁰The derivatives with respect to the charge density are always defined a part an additive constant, since $\int K \delta n(\mathbf{r}) d\mathbf{r} = 0$; this can be absorbed into the chemical potential μ . Otherwise, see Sec. 1.3.1, and Refs. [27] and [34] for the characterization of the domain for the variations of the charge density.

The two approaches of minimizing directly the Hohenberg-Kohn density functional (1.11) or solving for the associated differential equation (1.26) are equivalent¹¹, although they might differ in the computational effort associated with their implementation.

With these results in mind, the Thomas-Fermi method can be reinterpreted as a tentative approximation for the unknown universal functional $F[n(\mathbf{r})]$, taken to be the sum of the classical part of the electron-electron interaction plus the kinetic energy of the non-interacting electron gas in the local density approximation. This approach and its subsequent refinements, although very economical, have shown only moderate quantitative agreement with experimental data (see Ref. [37] for a review, and Ref. [21] for recent, more accurate, developments).

A much improved strategy has been developed by Kohn and Sham [4], that successfully readdressed the problem of finding a better approximation to the kinetic energy, albeit at the expense of having to deal with an additional set of auxiliary, orthonormal, orbitals. The kinetic energy is a one-body operator, and as such the expectation value (see Eq. 1.17) can be written in terms of the reduced density matrix of order 1:

$$T = -\frac{1}{2} \int \nabla^2 \gamma_1(\mathbf{r}', \mathbf{r}) \Big|_{\mathbf{r}'=\mathbf{r}} d\mathbf{r} = -\frac{1}{2} \sum_{i=1}^{\infty} w_i \langle \phi_i | \nabla^2 | \phi_i \rangle,$$

where the last equality comes from expressing the density matrix in the representation of its natural orbitals ϕ_i . In general, the natural orbitals are infinite in number. However, if N non-interacting electrons are considered, their ground state antisymmetric wavefunction can be expressed as the Slater

¹¹Note that (1.26) identifies the minimum only if there are no other extremal points; see Ref. [35] for a discussion and [36] for a proof that the *exact* F is indeed a convex functional.

determinant of only N orbitals, which are the lowest eigenstates of the non-interacting Hamiltonian.

The fundamental assumption of Kohn and Sham is to introduce a *reference system of non-interacting electrons in an external potential* $v_{KS}(\mathbf{r})$ such that the ground state charge density for this problem is the $n(\mathbf{r})$ that enters the Hohenberg-Kohn functional. This definition is meaningful (at least in its original formulation) if the charge density is non-interacting v -representable, so that there can be a unique definition for the total energy Kohn-Sham functional $T_s[n(\mathbf{r})] + \int v_{KS}(\mathbf{r})n(\mathbf{r}) d\mathbf{r}$, where T_s is now the kinetic energy associated with the new reference system. This technical problem can be overcome, again, by defining the kinetic energy part of the Kohn-Sham functional via a constrained search over all the Slater determinants of order N that integrate to a given charge density $n(\mathbf{r})$:

$$T_s[n(\mathbf{r})] = \min_{\Psi_S \rightarrow n} \langle \Psi_S | \hat{T}_c | \Psi_S \rangle = \min_{\Psi_S \rightarrow n} \left[\sum_{i=1}^N -\frac{1}{2} \int \psi_i^*(\mathbf{r}) \nabla^2 \psi_i(\mathbf{r}) d\mathbf{r} \right].$$

The $T_s[n(\mathbf{r})]$ can then be generalized by extending the search to all the antisymmetric wavefunctions that yield a charge density $n(\mathbf{r})$ (whether it is non-interacting v -representable or not) and it can be demonstrated [25] that this generalized T_s coincides with the original definition for all the densities that are non-interacting v -representable. With this definition at hand, a further decomposition of the universal functional F can take place via the introduction of the classical electrostatic energy term $E_H[n(\mathbf{r})]$:

$$F[n(\mathbf{r})] = T_s[n(\mathbf{r})] + E_H[n(\mathbf{r})] + E_{xc}[n(\mathbf{r})],$$

$$E_H[n(\mathbf{r})] = \frac{1}{2} \int \int \frac{n(\mathbf{r}_1)n(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2.$$

This decomposition, where F , T_s and the electrostatic energy are well-defined quantities, acts as an implicit definition for the *exchange-correlation energy* E_{xc} , that collects the contributions from the non-classic electrostatic interaction and from the difference between the true kinetic energy T and the non-interacting one T_s . The real success of the Kohn-Sham reformulation lies ultimately in the fact that E_{xc} , which is the term where all the inescapable complexity of the many-body problem has been finally pushed, is a small fraction of the total energy and, more importantly, that it can be *approximated* surprisingly well. These approximations, discussed in Chap. 2, are at present the strength of Density Functional Theory—in itself a very efficient reformulation for the quantum-mechanical problem—and its limit, in the very good but not exact approximation they provide.

The Euler-Lagrange equation (1.26) is rewritten with the new terminology as

$$\frac{\delta T_s[n(\mathbf{r})]}{\delta n(\mathbf{r})} + v_{KS}(\mathbf{r}) = \mu, \quad (1.27)$$

where the effective potential is

$$v_{KS}(\mathbf{r}) \equiv v(\mathbf{r}) + \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + v_{xc}(\mathbf{r}); \quad v_{xc}(\mathbf{r}) \equiv \frac{\delta E_{xc}[n(\mathbf{r})]}{\delta n(\mathbf{r})}$$

and, again, it is such that the set of non-interacting electrons with kinetic energy T_s in its field have the same ground-state charge density as the interacting electrons in the external potential v . The effective potential is now a function of the charge density itself: the problem has become self-consistent, and the solution for the reference system of non-interacting N -electrons

$$\Psi_S = \frac{1}{\sqrt{N!}} \det[\psi_1, \dots, \psi_N] \quad (1.28)$$

is now bound to produce a charge density that consistently outputs the original input effective potential.

$T_s[n(\mathbf{r})]$ as such is still an unknown functional of the density $n(\mathbf{r})$, but *now* it can be easily (but more expensively) written in terms of the N orbitals ψ_i for the non-interacting electrons, that act as auxiliary variables and properly span the domain of all the N -representable densities (provided that they are continuous and square integrable). The exact kinetic energy for the Kohn-Sham reference system is, straightforwardly:

$$T_s[n(\mathbf{r})] = -\frac{1}{2} \sum_{i=1}^N \langle \psi_i | \nabla^2 | \psi_i \rangle. \quad (1.29)$$

It should be noted that the ψ_i have to be orthonormal (solutions for the non-interacting fermion problem) in order for the kinetic-energy expression (1.29) to be valid.

The variational problem on $n(\mathbf{r})$ is thus finally reformulated in terms of a constrained search on the N ψ_i , where N^2 Lagrange multipliers λ_{ij} are introduced to provide for the orthonormality of the orbitals (and thus for the charge conservation):

$$\delta \left[T_s[n(\mathbf{r})] + \int v_{KS}(\mathbf{r})n(\mathbf{r}) d\mathbf{r} - \left(\sum_{i,j=1}^N \lambda_{ij} \langle \psi_j | \psi_i \rangle - \mathbb{I} \right) \right] = 0. \quad (1.30)$$

From this a single-particle Schrödinger-like equation is derived:

$$\left[-\frac{1}{2} \nabla^2 + v_{KS} \right] \psi_j = \sum_{i=1}^N \lambda_{ij} \psi_i; \quad (1.31)$$

λ_{ij} is a Hermitian matrix which can be diagonalized via a unitary transformation of the orbitals that leaves the non-interacting N -electron wavefunction (1.28) invariant, and consequently the density and the effective Hamiltonian

$-\frac{1}{2}\nabla^2 + v_{KS}$. From this, the “canonical” *Kohn-Sham equations* follow¹²:

$$\left[-\frac{1}{2}\nabla^2 + v_H(\mathbf{r}) + v_{xc}(\mathbf{r}) + v_{ext}(\mathbf{r}) \right] \psi_i(\mathbf{r}) = \hat{H}_{KS} \psi_i(\mathbf{r}) = \epsilon_i \psi_i(\mathbf{r}) \quad (1.32)$$

$$v_H(\mathbf{r}) = \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}', \quad v_{xc}(\mathbf{r}) = \frac{\delta E_{xc}}{\delta n(\mathbf{r})}$$

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

These equations should be iterated to self-consistency, until the minimum for the Kohn-Sham functional is reached; Kohn and Sham proposed taking the lowest N -orbitals to form the determinantal many-body wavefunction, and this choice works consistently for all the charge densities that are non-interacting v -representable (that is, for which T_s can be written in terms of single-particle orbitals).

The Kohn-Sham equations are the Euler-Lagrange equations for the constrained minimization of the Kohn-Sham functional. It should be noted that the search for the ground state can also proceed via the direct minimization of the full functional

$$E[\{\psi_i\}] = \sum_{i=1}^N -\frac{1}{2} \int \psi_i^*(\mathbf{r}) \nabla^2 \psi_i(\mathbf{r}) d\mathbf{r} + E_H[n(\mathbf{r})] + E_{xc}[n(\mathbf{r})] + \int v_{ext}(\mathbf{r}) n(\mathbf{r}) d\mathbf{r} \quad (1.33)$$

with respect to the N auxiliary orbitals ψ_i , with the proper constraints of orthonormality $\langle \psi_i | \psi_j \rangle = \delta_{ij}$ and charge density conservation. An alternative expression for the total energy can be derived by extracting the *band-energy term* $\sum_i \epsilon_i \equiv \text{tr} \left(\hat{T}_s + \hat{v}_{KS} \right)$:

$$E[\{\psi_i\}] = \sum_{i=1}^N \epsilon_i - E_H[n(\mathbf{r})] + E_{xc}[n(\mathbf{r})] - \int v_{xc}(\mathbf{r}) n(\mathbf{r}) d\mathbf{r}. \quad (1.34)$$

¹²A subscript is added for clarity to the external potential $v = v_{ext}$.

1.3.1 Fractional occupancies, Janak's theorem

An extension of the previous formalism that includes the case of fractional occupancies has been proposed by Janak [38]. Janak defines a generalized functional of $M (\geq N)$ orbitals $\{\psi_i\}$ and of their occupation numbers $\{f_i\}$; in the language of the constrained search formulation the functional is

$$T_J[n(\mathbf{r})] = \min_{\sum f_i |\psi_i|^2 \rightarrow n} \left[\sum_{i=1}^M f_i \int \psi_i^*(\mathbf{r}) \left(-\frac{1}{2} \nabla^2 \right) \psi_i(\mathbf{r}) d\mathbf{r} \right], \quad (1.35)$$

where the search is performed over all orthonormal orbitals and occupation numbers that yield the density $n(\mathbf{r})$ (hence $\sum_i f_i = N$). The condition $0 \leq f_i \leq 1 \forall i$ assures that the density matrix of order 1 is ensemble N -representable [32], and so it derives from an appropriate Γ_N ; consequently Eq. (1.22) provides the variational principle to develop a density-functional implementation in strict analogy with the Kohn-Sham discussion. The Janak's functional is formulated:

$$E[\{f_i\}, \{\psi_i\}] = \sum_{i=1}^M f_i \int \psi_i^*(\mathbf{r}) \left(-\frac{1}{2} \nabla^2 \right) \psi_i(\mathbf{r}) d\mathbf{r} + E_H[n(\mathbf{r})] + \\ + E_{J,xc}[n(\mathbf{r})] + \int v_{ext}(\mathbf{r}) n(\mathbf{r}) d\mathbf{r}, \quad (1.36)$$

where $E_{J,xc}$ is redefined to take into account the new definition for the non-interacting kinetic energy, and where now the minimization is performed with respect to the $\{\psi_i\}$ and the $\{f_i\}$, instead of the determinantal wave function (1.28) of the Kohn-Sham case. For a fixed set of $\{f_i\}$ the orbitals $\{\psi_i\}$ must satisfy the corresponding Euler-Lagrange equation:

$$\frac{\delta}{\delta \psi_i^*} \left\{ E[\{f_i\}, \{\psi_i\}] - \sum_{i=1}^M \lambda'_i (\langle \psi_i | \psi_i \rangle - 1) \right\} = 0, \quad \text{i.e.} \\ \left[-\frac{1}{2} f_i \nabla^2 + f_i \hat{v}_{J,KS} \right] \psi_i = \lambda'_i \psi_i, \quad (1.37)$$

where now the Lagrange multipliers are diagonal, and the orthogonality comes from the previous differential equation (1.37) (this is due to the choice of the T_J functional as a diagonal sum). These are again Kohn-Sham-like equations, and obviously for $f_i \neq 0$ $\epsilon_i = \lambda'_i/f_i$.

The dependence of the functional E on the occupation numbers can be seen by calculating the unconstrained variation of the energy with respect to one f_i , $\partial E/\partial f_i$, while allowing the orbitals to relax:

$$\begin{aligned} \frac{\partial E}{\partial f_i} &= -\frac{1}{2} \int \psi_i^*(\mathbf{r}) \nabla^2 \psi_i(\mathbf{r}) d\mathbf{r} + \int \frac{\delta(v_H + v_{xc} + v_{ext})}{\delta n(\mathbf{r})} \frac{\delta n(\mathbf{r})}{\delta f_i} d\mathbf{r} = \\ &= -\frac{1}{2} \int \psi_i^*(\mathbf{r}) \nabla^2 \psi_i(\mathbf{r}) d\mathbf{r} + \int v_{J,KS} \psi_i^*(\mathbf{r}) \psi_i(\mathbf{r}) d\mathbf{r} = \epsilon_i \end{aligned} \quad (1.38)$$

The expression above is referred to as *Janak's theorem*. The constrained variation of the energy functional with respect to f_i , that is the unconstrained minimization of $E - \mu N$, gives [39]:

$$\delta[E - \mu N] = \sum_i (\epsilon_i - \mu) \delta f_i, \quad (1.39)$$

which provides a rationale to the choice of $f_i = 0$ or 1 for ϵ_i greater or smaller than μ (so that $\delta[E - \mu N] \geq 0$). It is apparent from Eq. (1.39) that Janak's functional is not variational with respect to the occupation numbers (to a variation in the occupation numbers, precisely), as noted in [39][40].

Finally, it should be mentioned that there are some conceptual problems in the foundations of Eq. 1.38, as noted in Ref. [41], Sec. 9, and more recently in Ref. [34]. The objections arise in considering occupation numbers different from either 0 or 1 in zero-temperature density-functional theory, and in fact there are indications [27][34] that the domain of definition for the functional derivatives with respect to the charge density might not include

charge densities written in terms of fractional occupation numbers. Either on these methodological grounds, or following Janak's analysis, the conclusion is that occupation numbers (with an exception made for orbitals which will eventually be degenerate at the Fermi level) are bound to assume *at self-consistency* a value that is either 0 or 1. Fractional occupation numbers should thus be properly considered only in a statistical sense, as the output of a finite-temperature formulation.

1.3.2 Finite-temperature extension

The extension of density functional theory to the canonical ensemble has been discussed in Sec. 1.2.5, where the variational principle for the canonical Mermin functional $A_v[n(\mathbf{r})]$ has been demonstrated, following from Eq. 1.22. A Kohn-Sham mapping onto single-particle orbitals in a self-consistent potential will be discussed here, first for the case of non-interacting electrons and then for interacting ones.

If the electrons are non-interacting, the allowed states are the solutions of the single-particle equations

$$\left[-\frac{1}{2}\nabla^2 + \hat{v}_{ext} \right] \psi_i = \epsilon_i \psi_i. \quad (1.40)$$

The occupation of each state is determined by the statistics (that is, ultimately, a prescription for counting the states) and by the temperature; the requirement that the entropy $k_B \ln \Omega(E) = \sum_{\{f_i\}} W[\{f_i\}]$ is maximized, together with the Fermi-Dirac statistics, brings in the familiar Fermi-Dirac distribution [42]:

$$f_i = f \left(\frac{\epsilon_i - \mu}{T} \right) = \frac{1}{1 + \exp(\frac{\epsilon_i - \mu}{T})}; \quad (1.41)$$

for this distribution the entropy $S = k_B \ln \Omega(E)$ is

$$S = -k_B \sum_i [f_i \ln f_i + (1 - f_i) \ln(1 - f_i)] \quad (1.42)$$

and the expectation value of the kinetic energy operator is:

$$T_s = \sum_i f_i \int \psi_i^*(\mathbf{r}) \left(-\frac{1}{2}\nabla^2 \right) \psi_i(\mathbf{r}) d\mathbf{r} = \sum_i f_i \epsilon_i - \int v_{ext}(\mathbf{r}) n(\mathbf{r}) d\mathbf{r}. \quad (1.43)$$

These expressions can be derived more formally with the help of second-quantization operators (see Ref. [33], Chap. 6).

The Kohn-Sham formulation for a system of interacting electrons (see Refs. [43][41][44]) follows the lines of the zero-temperature case: a reference system of non-interacting electrons is introduced, for which the ground state has the charge density $n(\mathbf{r})$ that enters in the canonical Mermin functional $A_v[n(\mathbf{r})] = F_\beta[n(\mathbf{r})] + \int v_{ext}(\mathbf{r})n(\mathbf{r}) d\mathbf{r}$. An exchange-correlation potential $E_{\beta,xc}[n(\mathbf{r})]$ is then implicitly defined by isolating in $F_\beta[n(\mathbf{r})]$ the non-interacting kinetic energy T and entropy S of the non-interacting system defined in (1.43) and (1.42). The minimization of the Mermin functional, with the constraint of charge normalization, leads to the canonical Mermin-Kohn-Sham equations:

$$\left[-\frac{1}{2}\nabla^2 + v_H(\mathbf{r}) + v_{\beta,xc}(\mathbf{r}) + v_{ext}(\mathbf{r}) \right] \psi_i(\mathbf{r}) = \epsilon_i \psi_i(\mathbf{r}) \quad (1.44)$$

$$v_H(\mathbf{r}) = \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}', \quad v_{\beta,xc}(\mathbf{r}) = \frac{\delta E_{\beta,xc}}{\delta n(\mathbf{r})}$$

$$\sum_i f \left(\frac{\epsilon_i - \mu}{T} \right) = N, \quad n(\mathbf{r}) = \sum_i f_i |\psi_i(\mathbf{r})|^2.$$

These equations are formally similar to the zero-temperature case, although now there is a Lagrange multiplier μ to constrain the sum of the occupations to the total number of electrons N , since the number of states that the finite-temperature brings in jumps from N to ∞ (although for all purposes one can neglect all the states higher than a threshold for which the occupancies are practically zero). Alternatively, one can minimize directly the canonical Mermin functional as a function of the $\{f_i\}$ and $\{\psi_i\}$, with the explicit constraints of charge normalization *and* orthonormality:

$$A[T; \{\psi_i\}, \{f_i\}] = \sum_i f_i \int \psi_i^*(\mathbf{r}) \left(-\frac{1}{2}\nabla^2 \right) \psi_i(\mathbf{r}) d\mathbf{r} + E_H[n(\mathbf{r})] + E_{\beta,xc}[n(\mathbf{r})] + \int v_{ext}(\mathbf{r})n(\mathbf{r}) d\mathbf{r} + T \left(k_B \sum_i [f_i \ln f_i + (1 - f_i) \ln(1 - f_i)] \right). \quad (1.45)$$

An alternative expression can be obtained introducing the band-energy term:

$$A [T; \{\psi_i\}, \{f_i\}] = \sum_i f_i \epsilon_i - E_H[n(\mathbf{r})] + E_{\beta,xc}[n(\mathbf{r})] + \int v_{\beta,xc}(\mathbf{r})n(\mathbf{r}) d\mathbf{r} + T \left(k_B \sum_i [f_i \ln f_i + (1 - f_i) \ln(1 - f_i)] \right) \quad (1.46)$$

A different and very interesting approach has been proposed [45] (see also [46]), in which the kinetic-energy is not determined via a Kohn-Sham based orbital representation, but it is propagated a-la Feynman from high temperatures. Starting from the grand-canonical potential for the non-interacting electrons $\Omega = -\frac{1}{\beta} \ln \det [\mathbb{I} + \exp(-\beta(\hat{H} - \mu))]$, where \hat{H} is the single-particle $\hat{T}_e + \hat{v}_{ext}$, a Kohn-Sham decomposition is performed. The minimization of

$$A[n(\mathbf{r}); T] = -\frac{1}{\beta} \ln \det [\mathbb{I} + \exp(-\beta(\hat{H} - \mu))] + -E_H[n(\mathbf{r})] + E_{\beta,xc}[n(\mathbf{r})] - \int v_{\beta,xc}(\mathbf{r})n(\mathbf{r}) d\mathbf{r} \quad (1.47)$$

is then reached by iterating to self-consistency with the constraint of charge density normalization; the density matrix is evaluated as a path integral

$$e^{-\beta\hat{H}} = \left(e^{-\beta\hat{H}/P} \right)^P,$$

employing the Trotter decomposition, and with the choice of a real-space propagator for both the external potential *and* the kinetic energy operator. With this real-space formulation, this approach does not suffer from errors in the Brillouin Zone sampling of the kinetic-energy or of the charge density, although it does require a finer representation in real-space. In the present case, the minimization is reached iteratively, and not in a strictly variational fashion. In order to gauge the relative advantages and drawbacks a set of applications to a wider class of systems is needed; the relevant issues are

the relative costs of an orbital-based formulation in comparison to a charge-density one, the cutoff requirements for the chosen representation (e.g. plane waves vs. a real space grid), and the added computational stability that a strictly variational formulation does provide.