

<https://helda.helsinki.fi>

An Overview of Self-Consistent Field Calculations Within Finite Basis Sets

Lehtola, Susi

Multidisciplinary Digital Publishing Institute
2020-03-08

Lehtola, S.; Blockhuys, F.; Van Alsenoy, C. An Overview of Self-Consistent Field Calculations Within Finite Basis Sets. *Molecules* 2020, 25, 1218.

<http://hdl.handle.net/10138/348686>

Downloaded from Helda, University of Helsinki institutional repository.

This is an electronic reprint of the original article.

This reprint may differ from the original in pagination and typographic detail.

Please cite the original version.

Article

An Overview of Self-Consistent Field Calculations Within Finite Basis Sets [†]

Susi Lehtola ^{1,*} , Frank Blockhuys ² and Christian Van Alsenoy ^{2,*}¹ Department of Chemistry, University of Helsinki, P.O. Box 55 (A. I. Virtasen aukio 1) Helsinki, Finland² Department of Chemistry, University of Antwerp, Groenenborgerlaan 171, 2020 Antwerpen, Belgium; frank.blockhuys@uantwerpen.be

* Correspondence: susi.lehtola@alumni.helsinki.fi (S.L.); kris.vanalsenoy@uantwerpen.be (C.V.A.)

[†] The authors dedicate this paper to Prof. Paul Geerlings of the Free University of Brussels-VUB on the occasion of his 70th birthday.

Academic Editor: Ricardo Mosquera

Received: 26 December 2019; Accepted: 17 February 2020; Published: 8 March 2020



Abstract: A uniform derivation of the self-consistent field equations in a finite basis set is presented. Both restricted and unrestricted Hartree–Fock (HF) theory as well as various density functional approximations are considered. The unitary invariance of the HF and density functional models is discussed, paving the way for the use of localized molecular orbitals. The self-consistent field equations are derived in a non-orthogonal basis set, and their solution is discussed also in the presence of linear dependencies in the basis. It is argued why iterative diagonalization of the Kohn–Sham–Fock matrix leads to the minimization of the total energy. Alternative methods for the solution of the self-consistent field equations via direct minimization as well as stability analysis are briefly discussed. Explicit expressions are given for the contributions to the Kohn–Sham–Fock matrix up to meta-GGA functionals. Range-separated hybrids and non-local correlation functionals are summarily reviewed.

Keywords: self-consistent field theory; hartree-fock; density functional theory

1. Introduction

Electronic structure calculations have become a cornerstone of modern-day research in chemistry and materials physics, allowing *in silico* modeling of chemical reactions and the first principles design of novel catalysts [1]. Electronic structure calculations on molecular systems most often employ the linear combination of atomic orbitals (LCAO) approach, where the molecular orbitals (MOs) are expanded in terms of atomic orbitals (AOs). Several possible alternatives for the form of the AOs are commonly used—Gaussian-type orbitals (GTOs), Slater-type orbitals (STOs), as well as numerical atomic orbitals (NAOs); see [2] for details. LCAO electronic structure calculations involve a variational minimization of the total energy with respect to the AO expansion coefficients of the MOs. Importantly, the formalism used in the LCAO approach is not restricted to AOs which are atom-centered basis functions; it can also be used, e.g., in combination with fully numerical basis functions as in the finite element approach, as has been recently demonstrated in [3,4]. Once the energy has been minimized and the corresponding wave function has been obtained, it is possible to compute a number of properties either directly from the electronic wave function (e.g., electron densities, orbital energies, molecular dipole moment), or from its response to external perturbations (nuclear magnetic shieldings, vibrational frequencies, etc.).

The mathematical foundations for spin-restricted Hartree–Fock (HF) theory within the LCAO approach were laid out independently by Roothaan and Hall [5,6]. In their seminal papers, Roothaan and

Hall derived matrix equations that can be conveniently implemented on a computer as an iterative procedure. As will be seen later in Section 6, the Roothaan–Hall equations turn out to yield a generalized eigenvalue problem $\mathbf{FC} = \mathbf{SCE}$ in the non-orthogonal AO basis set, which had been solved some years before by Löwdin in the context of Heitler–London theory [7].

Subsequently to the work by Roothaan and Hall, Pople and Nesbet [8] and Berthier [9] independently published the corresponding equations for an unrestricted (open-shell) HF description by an analogous scheme, without providing an explicit derivation. The Pople–Nesbet–Berthier equations assume a form similar to the Roothaan–Hall equations—constituting a coupled set of general eigenvalue equations—as will also be seen later on in the manuscript (Section 6). Restricted open-shell HF was then described by Roothaan [10]; restricted open-shell calculations will not be considered in the present work as they have been extensively reviewed by Krebs in [11] to which we refer for further details.

Density functional theory [12,13] (DFT; see also [14,15]) became popular in chemistry through the efforts of Pople and coworkers in making the method generally available to quantum chemists [16] and showing that atomization energies from DFT may agree well with experiment [17,18]. Also DFT turns out to yield self-consistent field (SCF) equations that assume the same form as in HF but with a different expression for the Fock matrix \mathbf{F} . Pople and coworkers reported the equations necessary for solving SCF for DFT in the LCAO context up to generalized gradient approximation (GGA) functionals in [16]; an analogous derivation was also presented by Kobayashi et al in [19]. The self-consistent implementation of meta-GGA functionals was later described by Neumann, Nobes and Handy in [20]. Density functional calculations sometimes include also non-local correlation contributions; self-consistent LCAO implementations thereof have been reported by Vydrov and coworkers [21–24].

Despite the progress in and widespread success of DFT, to our knowledge, a uniform derivation of the SCF equations for HF and DFT including all the necessary expressions for the elements of the Kohn–Sham–Fock matrix up to the level of meta-GGA functionals has, up to now, not been explicitly published in the literature. This has likely contributed to the lack of complete support for meta-GGA functionals in popular quantum chemistry programs; for instance, PSI4 [25] and PySCF [26] lack support for meta-GGAs that depend on the Laplacian of the density such as the Becke–Roussel exchange functional [27], for example. This paper, therefore, presents such a derivation, yielding expressions of the DFT contributions to the Kohn–Sham–Fock matrix up to the level of meta-GGA functionals in a consistent way, facilitating the implementation of DFT in new programs.

The present derivation also has an obvious educational value. Indeed, in what follows, HF and various flavors of DFT belonging to different rungs of Jacob’s Ladder [28]—the local spin density approximation (LDA), the GGA and meta-GGA approximations—will be explicitly described in a uniform notation, making the similarities and dissimilarities between the approaches crystal clear. Facilitated by the uniform derivation, we will discuss key issues and features in the HF and DFT methodologies that arise from the mathematical formulation.

First, the basis set expansion of the molecular orbitals and the electron density is written out in Section 2. Then, the energy expression for HF and DFT is presented in Section 3, with a brief explanation of their physical content. The HF and DFT energy is shown to be invariant to rotations of the occupied and of the virtual orbitals in Section 4, allowing the construction of localized orbitals. The possibilities and drawbacks of spin-restricted calculations are discussed in Section 5. The finite-basis SCF equations are derived as generalized eigenvalue equations in Section 6. It is shown that the general eigenvalue equations can be reduced into normal eigenvalue equations by a transformation to an orthonormal basis in Section 7, and that linear dependencies in the basis can be eliminated on the way. The reason why the solution of the SCF equations amounts to a minimization of the total energy is rationalized in Section 8. Direct minimization methods are briefly introduced and stability analysis discussed in Section 9. The SCF method and direct minimization are contrasted in Section 10. Finally, the contributions to the Kohn–Sham–Fock

matrix arising from various-rung DFT functionals are listed in Section 11. The article concludes with a brief summary and discussion in Section 12. Atomic units are used throughout the text.

2. Basis Set Expansion

In the HF and DFT approaches, the electronic wave function is written as a Slater determinant, in which the electrons occupy a set of MOs $\varphi(\vec{r})$. The MOs are expanded in terms of normalized expansion functions $\chi(\vec{r})$, which are typically AOs. Both φ and χ , as well as the LCAO coefficients $C_{\mu i}$, are typically chosen to be real in the lack of magnetic interactions that would generally make the Hamiltonian operator complex. Note, however, that the use of complex coefficients has been shown to be sometimes beneficial to describe challenging systems even in the lack of magnetic fields in Hartree–Fock (as reviewed in [29]) or DFT (as shown in [30]); complex instabilities may also arise in specialized methods beyond the SCF level, see [31] for instance. The expansion functions are generally not orthonormal

$$\int d\vec{r} \chi_{\mu}(\vec{r})\chi_{\nu}(\vec{r}) = S_{\mu\nu} \neq \delta_{\mu\nu}$$

where $\delta_{\mu\nu}$ is the Kronecker delta: $\delta_{\mu\nu} = 1$ if $\mu = \nu$ and $\delta_{\mu\nu} = 0$ otherwise. Greek letters, $\mu, \nu, \lambda, \sigma, \eta, \zeta$ and θ will be used to identify the expansion functions $\chi(\vec{r})$, whereas Roman letters, i, j and k will be used to identify the MOs φ . The α (spin-up) and β (spin-down) MOs are expanded separately as

$$\varphi_i^{\alpha}(\vec{r}) = \sum_{\mu=1}^M C_{\mu i}^{\alpha} \chi_{\mu}(\vec{r}) \quad (1)$$

$$\varphi_i^{\beta}(\vec{r}) = \sum_{\mu=1}^M C_{\mu i}^{\beta} \chi_{\mu}(\vec{r}) \quad (2)$$

Both the α and β MOs are orthonormal to themselves

$$\int d\vec{r} \varphi_i^{\alpha}(\vec{r})\varphi_j^{\alpha}(\vec{r}) = \delta_{ij} \quad \text{and} \quad \int d\vec{r} \varphi_i^{\beta}(\vec{r})\varphi_j^{\beta}(\vec{r}) = \delta_{ij}$$

or equivalently within the basis set

$$(\mathbf{C}^{\alpha})^T \mathbf{S} \mathbf{C}^{\alpha} = \mathbf{1} \quad \text{and} \quad (\mathbf{C}^{\beta})^T \mathbf{S} \mathbf{C}^{\beta} = \mathbf{1}$$

However, the α orbitals are generally not orthonormal to the β orbitals:

$$\int d\vec{r} \varphi_i^{\alpha}(\vec{r})\varphi_j^{\beta}(\vec{r}) \neq \delta_{ij}$$

or

$$(\mathbf{C}^{\alpha})^T \mathbf{S} \mathbf{C}^{\beta} \neq \mathbf{1} \quad (3)$$

The electron density plays a pivotal role in quantum chemistry. In line with chemistry literature, $\rho(\vec{r})$ will be used to denote the electron density at the point \vec{r} in contrast to the physics notation $n(\vec{r})$ which is customary in the DFT literature. The total electron density is formed from the α and β densities, ρ_{α} and ρ_{β} , as $\rho(\vec{r}) = \rho_{\alpha}(\vec{r}) + \rho_{\beta}(\vec{r})$. The spin- σ electron density can be evaluated as

$$\rho_{\sigma}(\vec{r}) = \sum_{i=1}^{N_{\sigma}} |\varphi_i^{\sigma}(\vec{r})|^2 = \sum_{i=1}^{N_{\sigma}} \sum_{\mu\nu} C_{\mu i}^{\sigma} C_{\nu i}^{\sigma} \chi_{\mu}(\vec{r})\chi_{\nu}(\vec{r}) = \sum_{\mu\nu} P_{\mu\nu}^{\sigma} \chi_{\mu}(\vec{r})\chi_{\nu}(\vec{r}) \quad (4)$$

in which N_σ is the number spin- σ electrons in the system, and the sums over basis functions \sum_μ and \sum_ν indicate $\sum_{\mu=1}^M$ and $\sum_{\nu=1}^M$, respectively; this convention for easier readability of sums over basis functions is used throughout the rest of this work.

The density matrix \mathbf{P}^σ has been defined in Equation (4) as

$$P_{\mu\nu}^\sigma = \sum_{i=1}^{N_\sigma} C_{\mu i}^\sigma C_{\nu i}^\sigma \quad (5)$$

As is evident from the form of Equation (5), the density matrices are symmetric, $P_{\mu\nu}^\sigma = P_{\nu\mu}^\sigma$. As was already mentioned above, the total electron density is obtained from the sum of the α and β densities. Correspondingly, a total density matrix is given by

$$\mathbf{P} = \mathbf{P}^\alpha + \mathbf{P}^\beta \quad (6)$$

from which the total density can be evaluated using a relation analogous to Equation (4).

3. Energy Expression

The starting point for the derivation is the non-relativistic energy expression [5,8,13,16],

$$E = \sum_{\mu\nu} P_{\mu\nu} H_{\mu\nu} + \frac{1}{2} \sum_{\mu\nu\lambda\sigma} P_{\mu\nu} P_{\lambda\sigma} (\mu\nu|\lambda\sigma) - \frac{a}{2} \sum_{\mu\nu\lambda\sigma} (P_{\mu\lambda}^\alpha P_{\nu\sigma}^\alpha + P_{\mu\lambda}^\beta P_{\nu\sigma}^\beta) (\mu\nu|\lambda\sigma) + b \int f(\vec{r}) d\vec{r} \quad (7)$$

where the electron repulsion integral $(\mu\nu|\lambda\sigma)$ is defined as

$$(\mu\nu|\lambda\sigma) = \int d\vec{r}_1 \int d\vec{r}_2 \chi_\mu(\vec{r}_1) \chi_\nu(\vec{r}_1) \frac{1}{r_{12}} \chi_\lambda(\vec{r}_2) \chi_\sigma(\vec{r}_2) \quad (8)$$

and a and b are constants that define the fraction of HF exchange and the weight of the density functional approximation, respectively. The choice $a = 1$ and $b = 0$ corresponds to HF, whereas $a = 0$ and $b = 1$ yields a “pure” density functional without exact exchange such as the Perdew–Burke–Ernzerhof functional [32]. The choice $a \neq 0$ and $b \neq 0$ is the most general one, which corresponds to a hybrid functional [33] that are popular in quantum chemistry; perhaps the most famous example being the historical B3LYP functional [34].

The first term in Equation (7), which will be referred to as E^H , describes the kinetic energy of the electrons and the Coulombic attraction of the N nuclei in the system, with the matrix elements

$$H_{\mu\nu} = \int dr \chi_\mu(\vec{r}) \left(-\frac{1}{2} \nabla^2 + \sum_N \frac{Z_N}{r_N} \right) \chi_\nu(\vec{r}) \quad (9)$$

The one-electron operator in Equation (9) is commonly known as the core Hamiltonian, and the resulting E^H is the dominating contribution to the total energy.

However, the core Hamiltonian lacks electronic interactions. These are described by the second and third terms in Equation (7), which describe the classical Coulomb and the quantum mechanical “exchange” energy, and are referred to as E^J and E^K , respectively. The E^J contribution to the total energy can be straightforwardly derived from the expression for the Coulomb repulsion between the electrons described by the electron density $\rho(\vec{r})$

$$E^J = \frac{1}{2} \int d\vec{r}_1 \int d\vec{r}_2 \rho(\vec{r}_1) \frac{1}{r_{12}} \rho(\vec{r}_2)$$

whereas the expression for the exchange energy contribution E^K can be obtained, for instance, using Slater's rules for a HF wave function ($a = 1$).

The final term in Equation (7), referred to as E^{XC} , describes the DFT exchange-correlation contributions which, alike E^J and E^K , arise from electronic interactions. The exchange-correlation term is commonly written as

$$E^{XC} = \int d\vec{r} f(\vec{r}) = \int d\vec{r} \rho(\vec{r}) \epsilon^{XC}(\vec{r})$$

where ϵ^{XC} is the exchange-correlation energy density per electron. Usually $f(\vec{r})$ is a function of the electron density $\rho(\vec{r})$; it may also depend on the derivatives of $\rho(\vec{r})$ and the kinetic energy density $\tau(\vec{r})$, depending on which rung of Jacob's Ladder [28] is used to describe the exchange-correlation effects. The various rungs are discussed in Section 11.

4. Unitary Invariance

The \mathbf{P}^α and \mathbf{P}^β matrices turn out to be invariant to rotations of the occupied orbitals among themselves. Rotating the occupied subset of the molecular orbitals φ by a orthogonal matrix \mathbf{U} defines a new set of occupied orbitals

$$\varphi_i^{\alpha'} = \sum_{k=1}^{N_\alpha} \varphi_k^\alpha U_{ki}$$

the MO coefficients of which can be obtained as

$$C_{\mu i}^{\alpha'} = \sum_{k=1}^{N_\alpha} C_{\mu k}^\alpha U_{ki}$$

This can also be written in matrix notation as

$$\mathbf{C}^{\alpha'} = \mathbf{C}^\alpha \mathbf{U} \quad \text{or} \quad \mathbf{C}^{\alpha'} \mathbf{U}^T = \mathbf{C}^\alpha \quad (10)$$

The invariance to rotations in the occupied-occupied block is easy to prove, as

$$P_{\mu\nu}^{\alpha'} = \sum_{i=1}^{N_\alpha} C_{\mu i}^{\alpha'} C_{\nu i}^{\alpha'} = \sum_{ikl=1}^{N_\alpha} U_{ik} C_{\mu k}^\alpha U_{il} C_{\nu l}^\alpha = \sum_{kl=1}^{N_\alpha} \delta_{kl} C_{\mu k}^\alpha C_{\nu l}^\alpha = \sum_{k=1}^{N_\alpha} C_{\mu k}^\alpha C_{\nu k}^\alpha = P_{\mu\nu}^\alpha \quad (11)$$

where we have used the orthogonality of \mathbf{U} , $\mathbf{U}^T \mathbf{U} = \mathbf{1} = \mathbf{U} \mathbf{U}^T$.

The invariance to rotations in the occupied-occupied block can be used to fashion localized orbitals, for instance using an unitary optimization procedure [35]. Although localized orbitals are not strictly speaking observables—due to which several localization criteria have been suggested in the literature [36–40]—they have been shown to offer an effective way to study chemical reactions with ab initio calculations [41–43].

In addition to the occupied orbitals, in general there are also a number of unoccupied orbitals, which are commonly known as virtual orbitals. The number of virtual orbitals in any given calculation depends on the size of the basis set: the bigger the basis is, the more virtual orbitals there are. Because the virtual orbitals do not enter into the density matrix, the HF and DFT energy expression, Equation (7), is also invariant to rotations in the virtual-virtual block, similarly allowing their localization. However, as will be seen below in Section 9, the energy can be changed by mixing virtual orbitals into the occupied orbitals [44]. This approach provides another way to optimize the orbitals directly with, e.g., a gradient descent method, such as the geometric direct minimization method described in [45].

5. Spin-Restriction vs. Unrestriction

The molecular orbitals are obtained from the requirement that they minimize the total energy according to Equation (7). However, one must first choose the used spin formalism. The general choice is to use different spatial orbitals for the α and β electrons, in which case a spin-unrestricted approach is obtained. The unrestricted approach is often used even in systems in which there are an equal number of alpha and beta electrons, $N_\alpha = N_\beta$: although the spin-restricted and unrestricted descriptions often reproduce matching results for such systems near the equilibrium, only the unrestricted formalism is able to break bonds in general. The reason for this is that when molecules are stretched past the Coulson–Fischer point [46], the optimal orbitals spontaneously break spin symmetry, which can only be described in the unrestricted formalism. At variance, in the spin-restricted case the electrons occupy a common set of $N = N_\alpha = N_\beta = (N_\alpha + N_\beta)/2$ spatial orbitals. The limitation of the spatial orbitals to be the same for both spins, $C_{\mu i}^\alpha = C_{\mu i}^\beta = C_{\mu i}^\beta$, yields less variational freedom, and prevents the correct dissociation of e.g., the H_2 molecule. As a flip side, the spin-restricted formalism affords computational savings over the unrestricted approach. The spin-restricted density matrices, Equation (5), reduce to

$$P_{\mu\nu}^\alpha = P_{\mu\nu}^\beta = \frac{1}{2}P_{\mu\nu} = \sum_i^N C_{\mu i} C_{\nu i} \quad (12)$$

meaning, e.g., that the α and β exchange terms in Equation (7) coincide and can be simplified.

Spin-restriction is also possible in the case in which $N_\alpha \neq N_\beta$. In this case, a restricted open-shell method is obtained. Restricted open-shell methods are more involved than the spin-restricted and spin-unrestricted methods discussed in the present work. Restricted open-shell methods have been extensively discussed in [11] to which we refer for further discussion.

6. Self-Consistent Field Equations

Having chosen to use either spin-restricted or spin-unrestricted orbitals, one can proceed to the minimization of the energy expression in Equation (7). The energy expression depends only on the α and β density matrices \mathbf{P}^α and \mathbf{P}^β and their sum \mathbf{P} . The density matrices, in turn, are determined by the lowest N_α and N_β molecular orbitals according to Equation (5). Because the energy expression in Equation (7) thus only depends on the density matrices \mathbf{P}^α and \mathbf{P}^β , it is expedient to use the chain rule to write, e.g.,

$$\frac{\partial}{\partial C_{\theta k}^\alpha} = \sum_{\eta\zeta} \frac{\partial P_{\eta\zeta}^\alpha}{\partial C_{\theta k}^\alpha} \frac{\partial}{\partial P_{\eta\zeta}^\alpha} \quad (13)$$

where the partial derivative of the density matrix element $P_{\eta\zeta}^\alpha$ is

$$\begin{aligned} \frac{\partial P_{\eta\zeta}^\alpha}{\partial C_{\theta k}^\alpha} &= \frac{\partial}{\partial C_{\theta k}^\alpha} \sum_{i=1}^{N_\alpha} C_{\eta i}^\alpha C_{\zeta i}^\alpha \\ &= \sum_{i=1}^{N_\alpha} \delta_{\theta\eta} \delta_{ki} C_{\zeta i}^\alpha + \sum_{i=1}^{N_\alpha} C_{\eta i}^\alpha \delta_{\theta\zeta} \delta_{ki} \\ &= \delta_{\theta\eta} C_{\zeta k}^\alpha + C_{\eta k}^\alpha \delta_{\theta\zeta} \end{aligned} \quad (14)$$

The β orbital derivative of the total density matrix has the same form as Equation (14), where all α are replaced with β . Note that these findings hold even when the same orbitals are used for both α and β in a spin-restricted formalism, since the α and β orbitals are formally independent.

Due to the chain rule, Equation (13), all we need are the density matrix derivatives of the energy expression. We only have to calculate the derivatives of the energy expression of Equation (7) for one spin, as the energy expression is symmetric with respect to the α and β densities. It does not matter which spin we choose to be “up”; the expressions for the other spin will follow by symmetry by interchanging $\alpha \leftrightarrow \beta$. The first term of Equation (7) yields simply

$$\frac{\partial E^H}{\partial P_{\eta\zeta}^\alpha} = \frac{\partial}{\partial P_{\eta\zeta}^\alpha} \sum_{\mu\nu} P_{\mu\nu} H_{\mu\nu} = H_{\eta\zeta} \quad (15)$$

Next, taking the partial derivative with respect to $P_{\eta\zeta}^\alpha$ of the Coulomb and exchange terms in Equation (7) results in

$$\frac{\partial E^J}{\partial P_{\eta\zeta}^\alpha} = \sum_{\mu\nu} (P_{\mu\nu}^\alpha + P_{\mu\nu}^\beta) (\mu\nu|\eta\zeta) = \sum_{\mu\nu} P_{\mu\nu} (\mu\nu|\eta\zeta) = J_{\eta\zeta} \quad (16)$$

where J is known as the Coulomb matrix, and

$$\frac{\partial E^K}{\partial P_{\eta\zeta}^\alpha} = -a \sum_{\mu\nu} P_{\mu\nu}^\alpha (\mu\eta|v\zeta) = -a K_{\eta\zeta}^\alpha \quad (17)$$

where K^α is the spin- α exchange matrix, respectively. The Coulomb and exchange matrices can be used to rewrite the energy expression in Equation (7) as

$$E = \sum_{\mu\nu} P_{\mu\nu} H_{\mu\nu} + \frac{1}{2} \sum_{\mu\nu} P_{\mu\nu} J_{\mu\nu} - \frac{a}{2} \sum_{\mu\nu} (P_{\mu\nu}^\alpha K_{\mu\nu}^\alpha + P_{\mu\nu}^\beta K_{\mu\nu}^\beta) + b \int f(\vec{r}) d\vec{r} \quad (18)$$

Note that in contrast to the Coulomb and exact exchange terms, the exchange-correlation term does not undergo simplifications, because the exchange-correlation term is not quadratic in the density matrix, as will be seen later in Section 11. For the time being, we will denote the partial derivative of E^{XC} with respect to $P_{\eta\zeta}^\alpha$ as

$$\frac{\partial E^{XC}}{\partial P_{\eta\zeta}^\alpha} = b K_{\eta\zeta}^{XC;\alpha} \quad (19)$$

as the full expressions for $K^{XC;\alpha}$ will be presented in Section 11. Now, collecting the partial derivatives in Equations (15)–(19) gives us the density matrix derivatives of the energy expression as

$$\frac{\partial E}{\partial P_{\eta\zeta}^\sigma} = H_{\eta\zeta} + J_{\eta\zeta} - a K_{\eta\zeta}^\sigma + b K_{\eta\zeta}^{XC;\sigma} = F_{\eta\zeta}^\sigma \quad (20)$$

where we have identified the Kohn–Sham–Fock matrices F^σ , where σ denotes α or β . Because the density matrices defined by Equation (5) are symmetric, also the Fock matrices are symmetric, $F_{\eta\zeta}^\sigma = F_{\zeta\eta}^\sigma$. Note that since the Fock matrices only depend on the density matrices, also they are invariant to occupied-occupied and virtual-virtual rotations, $F^{\sigma'} = F^\sigma$.

Naïvely, one would obtain the orbital derivative of the full energy expression in Equation (7) with Equations (13), (14) and (20) and then set it to zero to yield an equation for the unknown expansion coefficients C^α . However, the molecular orbitals cannot be varied freely: one must make sure that the

orbitals stay orthonormal during the variation, as otherwise the Pauli exclusion principle would be violated. For instance, the orthonormality condition for the α electrons is

$$\int \varphi_i^\alpha(\vec{r}) \varphi_j^\alpha(\vec{r}) d\vec{r} = \delta_{ij} \quad (21)$$

The way to enforce these conditions is to use Lagrangian multipliers ε_{ij} . That is, instead of the bare energy expression E , we will optimize the Lagrangian

$$L = E - \sum_{ij} \varepsilon_{ij}^\alpha \left[\int d\vec{r} \varphi_i^\alpha(\vec{r}) \varphi_j^\alpha(\vec{r}) - \delta_{ij} \right] - \sum_{ij} \varepsilon_{ij}^\beta \left[\int d\vec{r} \varphi_i^\beta(\vec{r}) \varphi_j^\beta(\vec{r}) - \delta_{ij} \right] \quad (22)$$

where the sums over i and j run over all orbitals; that is, both the occupied and the virtual ones. We can see from Equation (22) that the matrices of Lagrangian multipliers ε^α and ε^β can be chosen to be symmetric. For instance, if ε^α contained a symmetric part ε_s^α and an antisymmetric part ε_a^α , $\varepsilon^\alpha = \varepsilon_s^\alpha + \varepsilon_a^\alpha$, the contribution from the antisymmetric part would vanish because it is multiplied with the orbital overlap that is symmetric.

Next, we can calculate $\partial L / \partial C_{\theta k}^\alpha$, where $\partial E / \partial C_{\theta k}^\alpha$ is given by Equations (13), (14), (20) and the derivative of the constraint term is given by

$$\begin{aligned} \frac{\partial}{\partial C_{\theta k}^\alpha} \sum_{ij} \varepsilon_{ij}^\alpha \left[\int d\vec{r} \varphi_i^\alpha(\vec{r}) \varphi_j^\alpha(\vec{r}) - \delta_{ij} \right] &= \sum_{ij} \varepsilon_{ij}^\alpha \frac{\partial}{\partial C_{\theta k}^\alpha} \sum_{\eta \zeta} \left[C_{\eta i}^\alpha S_{\eta \zeta} C_{\zeta j}^\alpha - \delta_{ij} \right] \\ &= \sum_{ij} \varepsilon_{ij}^\alpha \sum_{\eta \zeta} \left[\delta_{\eta \theta} \delta_{k i} C_{\zeta j}^\alpha + C_{\eta i}^\alpha \delta_{\zeta \theta} \delta_{k j} \right] S_{\eta \zeta} \\ &= \sum_{\eta} \sum_i \varepsilon_{ki}^\alpha C_{\eta i}^\alpha S_{\theta \eta} + \sum_{\eta} \sum_i \varepsilon_{ik}^\alpha C_{\eta i}^\alpha S_{\eta \theta} \end{aligned} \quad (23)$$

where on the third line dummy summation indices have been renamed from j to i and ζ to η . The derivative can be evaluated as

$$\begin{aligned} \frac{\partial L}{\partial C_{\theta k}^\alpha} &= \sum_{\eta \zeta} \frac{\partial E}{\partial P_{\eta \zeta}^\alpha} \frac{\partial P_{\eta \zeta}^\alpha}{\partial C_{\theta k}^\alpha} - \sum_{ij} \varepsilon_{ij}^\alpha \frac{\partial}{\partial C_{\theta k}^\alpha} \int d\vec{r} \varphi_i^\alpha(\vec{r}) \varphi_j^\alpha(\vec{r}) \\ &= \sum_{\eta \zeta} F_{\eta \zeta}^\alpha \left[\delta_{\theta \eta} C_{\zeta k}^\alpha + C_{\eta k}^\alpha \delta_{\theta \zeta} \right] - \sum_{\zeta} \sum_i \varepsilon_{ki}^\alpha C_{\zeta i}^\alpha S_{\theta \zeta} - \sum_{\eta} \sum_i \varepsilon_{ik}^\alpha C_{\eta i}^\alpha S_{\eta \theta} \\ &= 2 \sum_{\eta} F_{\theta \eta}^\alpha C_{\eta k}^\alpha - 2 \sum_{\eta} \sum_i \varepsilon_{ki}^\alpha S_{\theta \eta} C_{\eta i}^\alpha \end{aligned} \quad (24)$$

because \mathbf{F} and ε are symmetric, and dummy summation indices can be renamed.

The optimal orbitals satisfy the stationary condition $\partial L / \partial C_{\theta k}^\alpha = 0$ from which

$$\sum_{\eta} F_{\theta \eta}^\alpha C_{\eta k}^\alpha = \sum_{\eta} \sum_i S_{\theta \eta} C_{\eta i}^\alpha \varepsilon_{ik}^\alpha \quad (25)$$

Equation (25) can thus be written in matrix form as

$$\mathbf{F}^\alpha \mathbf{C}^\alpha = \mathbf{S} \mathbf{C}^\alpha \mathbf{E}^\alpha \quad (26)$$

where \mathbf{E}^α is the (symmetric) matrix of Lagrangian multipliers.

Because \mathbf{E}^α is symmetric, it can be diagonalized and it has real eigenvalues. Let us now assume that \mathbf{U}^α is an orthogonal matrix that diagonalizes \mathbf{E}^α

$$E_{ij}^\alpha \rightarrow E_{ij}^{\alpha'} = \sum_{kl} U_{ki}^\alpha E_{kl}^\alpha U_{lj}^\alpha = \delta_{ij} \varepsilon_i^\alpha \quad (27)$$

where ε^α are the eigenvalues. Re-expressing the orbital coefficients \mathbf{C}^α in terms of a new set of orbitals rotated by \mathbf{U}^α with Equation (10), $\mathbf{C}^\alpha = \mathbf{C}^{\alpha'} (\mathbf{U}^\alpha)^\top$, rewrites Equation (26) in the form

$$\mathbf{F}^\alpha \mathbf{C}^{\alpha'} (\mathbf{U}^\alpha)^\top = \mathbf{S} \mathbf{C}^{\alpha'} (\mathbf{U}^\alpha)^\top \mathbf{E}^\alpha$$

that can be multiplied from the right by \mathbf{U}^α producing

$$\mathbf{F}^\alpha \mathbf{C}^{\alpha'} = \mathbf{S} \mathbf{C}^{\alpha'} \mathbf{E}^{\alpha'} \quad (28)$$

where, according to Equation (27), $\mathbf{E}^{\alpha'} = (\mathbf{U}^\alpha)^\top \mathbf{E}^\alpha \mathbf{U}^\alpha$ is a diagonal matrix with elements ε_i^α .

Equation (28) is almost what we want—an equation in the rotated basis that looks like Equation (26) with a diagonal $\mathbf{E}^{\alpha'}$ —but one problem remains: the Kohn–Sham–Fock matrix is still the one corresponding to the original orbitals \mathbf{C}^α instead of the transformed orbitals $\mathbf{C}^{\alpha'}$, while the orbital rotation by \mathbf{U} that takes us from \mathbf{C}^α to $\mathbf{C}^{\alpha'}$ may lead to a different Kohn–Sham–Fock matrix $\mathbf{F}^{\alpha'} \neq \mathbf{F}^\alpha$. However, if we choose the form of \mathbf{U} such that the occupied-virtual (ov) and virtual-occupied (vo) blocks vanish

$$\mathbf{U} = \begin{pmatrix} \mathbf{U}_{oo} & \mathbf{U}_{ov} \\ \mathbf{U}_{vo} & \mathbf{U}_{vv} \end{pmatrix} = \begin{pmatrix} \mathbf{U}_{oo} & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_{vv} \end{pmatrix} \quad (29)$$

then \mathbf{U} only rotates occupied orbitals with occupied orbitals and virtual orbitals with virtual orbitals, meaning that the orbital rotation does not change the density matrix given in Equation (11). Then, the Fock matrix corresponding to the rotated orbitals coincides with that of the original orbitals, $\mathbf{F}' = \mathbf{F}$, completing the proof that \mathbf{E}^α can be chosen to be diagonal. (Occupied-virtual rotations \mathbf{U}_{ov} or \mathbf{U}_{vo} , discussed in more detail in Section 9, are in fact here forbidden: the SCF equations were derived with the assumption that the energy is stationary, but this condition would instantly be violated by such rotations.)

We have thus obtained the Berthier–Pople–Nesbet [8,9,13,16] equations for the orbital coefficients

$$\begin{cases} \mathbf{F}^\alpha \mathbf{C}^\alpha = \mathbf{S} \mathbf{C}^\alpha \mathbf{E}^\alpha \\ \mathbf{F}^\beta \mathbf{C}^\beta = \mathbf{S} \mathbf{C}^\beta \mathbf{E}^\beta \end{cases} \quad (30)$$

where the primes have become unnecessary and have been omitted for simplicity. The elements of the Kohn–Sham–Fock matrices \mathbf{F}^α and \mathbf{F}^β are given by Equation (20), and the orbital energy matrices \mathbf{E}^α and \mathbf{E}^β are diagonal. In the spin-restricted case [5,16] the α and β molecular orbitals coincide, leading to identical density matrices \mathbf{P}^α and \mathbf{P}^β , and identical Fock matrices \mathbf{F}^α and \mathbf{F}^β . In this case, the SCF equations simplify to the Roothaan–Hall form

$$\mathbf{FC} = \mathbf{SCE} \quad (31)$$

which was already mentioned in the Introduction.

7. Solution of Self-Consistent Field Equations

The Roothaan–Hall and Berthier–Pople–Nesbet expressions take the form of a generalized eigenvalue equation. The conventional way to solve these equations is to re-express the (unknown) orbital coefficients in terms of a matrix \mathbf{X} as

$$\mathbf{C} = \mathbf{X}\tilde{\mathbf{C}} \quad (32)$$

Inserting Equation (32) into the Roothaan–Hall equation, Equation (31), yields

$$\mathbf{F}\mathbf{X}\tilde{\mathbf{C}} = \mathbf{S}\mathbf{X}\tilde{\mathbf{C}}\mathbf{E}$$

which can be multiplied from the left with \mathbf{X}^T to yield

$$\mathbf{X}^T\mathbf{F}\mathbf{X}\tilde{\mathbf{C}} = \mathbf{X}^T\mathbf{S}\mathbf{X}\tilde{\mathbf{C}}\mathbf{E}$$

This means that the orbital transform of Equation (32) yields a new generalized eigenvalue equation

$$\tilde{\mathbf{F}}\tilde{\mathbf{C}} = \tilde{\mathbf{S}}\tilde{\mathbf{C}}\mathbf{E} \quad (33)$$

where $\tilde{\mathbf{F}} = \mathbf{X}^T\mathbf{F}\mathbf{X}$ and $\tilde{\mathbf{S}} = \mathbf{X}^T\mathbf{S}\mathbf{X}$. Now, if we choose \mathbf{X} in such a way that $\tilde{\mathbf{S}} = \mathbf{1}$, Equation (33) reduces to a normal eigenvalue equation

$$\tilde{\mathbf{F}}\tilde{\mathbf{C}} = \tilde{\mathbf{C}}\mathbf{E} \quad (34)$$

which can be solved with standard techniques. Then, the wanted orbital coefficients \mathbf{C} can be calculated from $\tilde{\mathbf{C}}$ using Equation (32).

If the basis set is well-conditioned, the matrix \mathbf{X} can be chosen as

$$\mathbf{X} = \mathbf{V}\mathbf{\Lambda}^{-1/2}\mathbf{V}^T \quad (35)$$

where \mathbf{V} and $\mathbf{\Lambda}$ are the eigenvectors and eigenvalues of \mathbf{S}

$$\mathbf{S} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T \quad (36)$$

This procedure is known as symmetric orthogonalization [7].

However, if a large LCAO basis is used, the atomic orbital basis functions centered on different atoms may generate linear dependencies in the basis, making the basis set expansion ambiguous. These linear dependencies can be removed with the “canonical” orthonormalization procedure [47], in which

$$\mathbf{X} = \mathbf{V}'\mathbf{\Lambda}'^{-1/2} \quad (37)$$

where only those eigenvectors \mathbf{V}_i with large enough eigenvalues $\lambda_i \geq \tau$ are included. The threshold τ is typically of the order of $\tau = 10^{-7} \dots 10^{-5}$, and its value may have a noticeable effect on, e.g., the absolute energies that result from a SCF calculation; relative energies, however, should be less sensitive to τ . If no eigenvalues fall under the threshold τ , the symmetric and canonical orthogonalization approaches become equivalent for the purposes of SCF calculations in the case of a well-conditioned basis set: both yield an orthonormal basis of the same size, which will yield the same variational ground state energy.

Unnormalized basis sets can also be handled easily by the orthogonalization procedure. Although in principle it is not necessary to normalize the individual basis functions before obtaining an orthonormal basis by Equations (35) and (37), computer linear algebra packages may fail to find the eigenvalues and eigenvectors in a reliable fashion if the basis functions have pronouncedly different norms. Moreover,

missing normalization of the basis set affects the eigenvalues, which has repercussions for canonical orthogonalization. These issues can be circumvented by normalizing the overlap matrix $\mathbf{S} \rightarrow \mathbf{S}' = \mathbf{nS}$ where $n_{ij} = S_{ii}^{-1/2} \delta_{ij}$ before using Equations (35) and (37) [3,4]. The orthogonalizer for the unnormalized basis set is obtained as $\mathbf{X} \rightarrow \mathbf{nX}$; it is easy to see that this satisfies the necessary condition $\mathbf{X}^T \mathbf{S} \mathbf{X} = \mathbf{1}$ even though the symmetricity of \mathbf{X} for the case of Equation (35) will be lost.

Even if \mathbf{S} has been properly normalized, the use of the symmetric or canonical orthogonalization procedures still requires that the diagonalization of \mathbf{S} is numerically stable. However, whenever a large number of linear dependencies exists in the basis set (e.g., a large number of diffuse functions are used or two nuclei are close together), \mathbf{S} may become so ill-conditioned it cannot be accurately diagonalized. In such cases it is possible to reduce the size of the basis set without losing a significant amount of accuracy by an automatic procedure, see [48,49] for details.

8. Why Does the Self-Consistent Field Method Minimize the Energy?

The SCF equations, Equation (30) or Equation (31), offer a way to solve for the molecular orbitals described by \mathbf{C} from a Kohn–Sham–Fock matrix \mathbf{F} by finding its eigenvectors from Equation (34). However, the Kohn–Sham–Fock matrix depends on the density matrices, which are built from the molecular orbitals according to the Aufbau principle. In the SCF procedure, one tries to find a self-consistent solution: \mathbf{C} yields \mathbf{F} , whose eigenvectors are \mathbf{C} . The procedure starts from an initial guess for the orbitals \mathbf{C}^σ or the density matrices \mathbf{P}^σ , which have been recently reviewed and benchmarked in [50] to which we refer for further details.

Why does the self-consistent field procedure—diagonalizing \mathbf{F}^σ to update the orbital coefficients \mathbf{C}^σ —correspond to minimization of the Hartree–Fock/Kohn–Sham energy? For simplicity, let us examine the case of HF theory. The energy expression, Equation (18), can be written in this case ($a = 1, b = 0$) as

$$E = \text{Tr } \mathbf{P} \mathbf{H} + \frac{1}{2} \text{Tr } \mathbf{P} \mathbf{J} - \frac{1}{2} \text{Tr } \mathbf{P}^\alpha \mathbf{K}^\alpha - \frac{1}{2} \text{Tr } \mathbf{P}^\beta \mathbf{K}^\beta \quad (38)$$

The Fock matrix elements, Equation (20), are given by

$$\mathbf{F}^\alpha = \mathbf{H} + \mathbf{J} - \mathbf{K}^\alpha \quad (39)$$

$$\mathbf{F}^\beta = \mathbf{H} + \mathbf{J} - \mathbf{K}^\beta \quad (40)$$

Equation (38) can be rewritten with Equations (39) and (40) as

$$E = \frac{1}{2} \text{Tr } \mathbf{P}^\alpha (\mathbf{H} + \mathbf{F}^\alpha) + \frac{1}{2} \text{Tr } \mathbf{P}^\beta (\mathbf{H} + \mathbf{F}^\beta) \quad (41)$$

Expanding the density matrices using Equation (5) we see that Equation (41) can be written as

$$E = \frac{1}{2} \sum_{i=1}^{N_\alpha} (h_{ii}^\alpha + f_{ii}^\alpha) + \frac{1}{2} \sum_{i=1}^{N_\beta} (h_{ii}^\beta + f_{ii}^\beta) \quad (42)$$

where the core Hamiltonian and Fock matrices have been written in the molecular orbital basis, $\mathbf{h}^\sigma = (\mathbf{C}^\sigma)^T \mathbf{H} \mathbf{C}^\sigma$ and $\mathbf{f}^\sigma = (\mathbf{C}^\sigma)^T \mathbf{F}^\sigma \mathbf{C}^\sigma$.

If one were to start the calculation from the core guess, then $\sum_i h_{ii}^\alpha$ and $\sum_i h_{ii}^\beta$ would be minimized. However, as discussed in [50], this is a horrible choice as it completely disregards electronic repulsion effects, meaning that the $\sum_i f_{ii}^\alpha$ and $\sum_i f_{ii}^\beta$ terms are far from optimal. The Roothaan step—obtaining new molecular orbitals by diagonalization of the Fock matrix—results in a minimization of the $\sum_i f_{ii}^\alpha$ and $\sum_i f_{ii}^\beta$

terms, as after diagonalization only the lowest orbitals become populated and the sum thus runs only over the lowest eigenvalues f_{ii}^σ . After the update, $\sum_i h_{ii}^\alpha$ and $\sum_i h_{ii}^\beta$ no longer yield their lowest possible values. However, the increase in the value of $\sum_i h_{ii}^\alpha + \sum_i h_{ii}^\beta$ should be much smaller than the decrease in the value of $\sum_i f_{ii}^\alpha + \sum_i f_{ii}^\beta$, as the Fock matrices \mathbf{f}^α and \mathbf{f}^β also contain the core Hamiltonian. It is thus seen that Roothaan's self-consistent field method, that is, the iterative diagonalization of the Fock matrix minimizes the energy.

However, the minimization is only valid for a fixed potential \mathbf{f}^σ in which the electrons are moving. When the orbitals are changed—as happens when \mathbf{f} is made diagonal and its lowest eigenvectors occupied—a new Fock matrix \mathbf{F} must be built and a new \mathbf{f} constructed: the potential also changes with the electron density. If the orbitals were far from their optimal values, \mathbf{P} and therefore \mathbf{F} may change quite radically by the orbital update. This means that even though \mathbf{f} was made diagonal in the previous iteration, it is no longer diagonally dominant after it has been updated. Indeed, the straightforward iterative diagonalization procedure often fails to converge for all but the simplest systems, because the density tends to undergo large oscillations in the naïve self-consistency cycle. To make the method usable, the convergence of the fixed-point problem of finding a \mathbf{C} that generates \mathbf{F} that generates \mathbf{C} must be stabilized or accelerated in some way. This can be achieved, e.g., by damping [51,52], level shifts [53–55], or extrapolation [56–59]. Fractional occupations can also be used in the initial iterations to aid convergence [60].

The argument for why density functional calculations converge similarly to HF with the iterative Roothaan procedure is somewhat less obvious, because unlike HF the exchange-correlation functional is not generally quadratic in the density. However, the total energy expression is approximately quadratic also in DFT when one is sufficiently close to an extremal point, as is easily seen by a Taylor expansion of Equation (7). In practice the iterative procedure works well also for DFT, whose contributions to the Kohn–Sham–Fock matrix we will discuss in Section 11.

9. Direct Minimization of the Energy

Instead of solving the orbitals from the SCF equations, which were obtained in Section 6 from the stationary condition for the energy under the constraint of orthonormal orbitals, the orbitals can also be optimized by a direct minimization of the energy. As was discussed in Section 4, the energy expression of Equation (7) is invariant to occupied-occupied and virtual-virtual rotations. This means that if we have o^σ occupied orbitals and v^σ virtual orbitals from some initial guess (see possible choices in [50]) for spin σ , we can consider the energy as a function of a set of $o^\sigma v^\sigma$ rotation angles [44] by examining a rotation of the orbitals via Equation (10) by an orthogonal matrix

$$\mathbf{U}^\sigma(\boldsymbol{\theta}^\sigma) = \exp \begin{pmatrix} \mathbf{0} & \boldsymbol{\theta}^\sigma \\ -(\boldsymbol{\theta}^\sigma)^\top & \mathbf{0} \end{pmatrix} \quad (43)$$

where $\boldsymbol{\theta}^\sigma$ is an $o^\sigma \times v^\sigma$ matrix containing the rotation angles. The rotation matrix determined by Equation (43) reduces to an identity matrix for vanishing rotation parameters, $\boldsymbol{\theta}^\sigma = \mathbf{0}$. Because the rotation matrix of Equation (43) is orthogonal, it automatically preserves the orthonormality of the orbitals, and special tricks i.e. Lagrangian multipliers are not needed to enforce this behavior.

The change in the density matrix is given by

$$\frac{\partial P_{\eta\zeta}^\sigma}{\partial \theta_{ia}^\sigma} = \sum_{k=1}^{N_\sigma} \left[\frac{\partial C_{\eta k}^\sigma}{\partial \theta_{ia}^\sigma} C_{\zeta k}^\sigma + C_{\eta k}^\sigma \frac{\partial C_{\zeta k}^\sigma}{\partial \theta_{ia}^\sigma} \right] \quad (44)$$

How do the orbital coefficients change? Remembering that the first N_σ orbitals are occupied, and the rest are virtual, we can write $\mathbf{C}^\sigma = \begin{pmatrix} \mathbf{C}_o^\sigma & \mathbf{C}_v^\sigma \end{pmatrix}$. After an infinitesimal rotation $\boldsymbol{\theta} \approx \mathbf{0}$, the occupied orbitals change into $\mathbf{C}_o^{\sigma'} \approx \mathbf{C}_o^\sigma - \mathbf{C}_o^\sigma (\boldsymbol{\theta}^\sigma)^\top$, that is

$$C_{\eta k}^{\sigma'} = C_{\eta k}^\sigma - \sum_{b \text{ virtual}} C_{\eta b}^\sigma \theta_{kb}^\sigma \quad (45)$$

from which $\partial C_{\eta k}^\sigma / \partial \theta_{ia}^\sigma = -C_{\eta a}^\sigma \delta_{ik}$. Now the gradient of the energy with respect to rotation of the current set of orbitals can be obtained as

$$\frac{\partial E}{\partial \theta_{ia}^\sigma} = \sum_{\eta \zeta} \frac{\partial E}{\partial P_{\eta \zeta}^\sigma} \frac{\partial P_{\eta \zeta}^\sigma}{\partial \theta_{ia}^\sigma} = \sum_{\eta \zeta} F_{\eta \zeta}^\sigma \sum_{k=1}^{N_\sigma} \left[\frac{\partial C_{\eta k}^\sigma}{\partial \theta_{ia}^\sigma} C_{\zeta k}^\sigma + C_{\eta k}^\sigma \frac{\partial C_{\zeta k}^\sigma}{\partial \theta_{ia}^\sigma} \right] = - \sum_{\eta \zeta} F_{\eta \zeta}^\sigma \left[C_{\eta a}^\sigma C_{\zeta i}^\sigma + C_{\eta i}^\sigma C_{\zeta a}^\sigma \right] = -2f_{ia}^\sigma \quad (46)$$

where $\mathbf{f}^\sigma = (\mathbf{C}^\sigma)^\top \mathbf{F}^\sigma \mathbf{C}^\sigma$ is the Fock matrix in the MO basis. Direct minimization of Equation (7) can then be pursued using Equation (46) with, e.g., gradient descent methods. However, a proper preconditioning of the search direction is essential in order for the algorithm to be usable; see, e.g., the geometric direct minimization method described in [45]. Many other direct minimization methods for the HF or DFT energy have also been proposed, and we refer the interested reader to the vast existing literature that cannot be comprehensively cited here.

10. SCF vs. Direct Minimization

Having described two alternative ways for solving the orbitals, we can discuss their advantages and disadvantages. The self-consistent field method is hard to beat for systems where convergence is straightforward: a suitably stabilized and accelerated SCF procedure often converges within 10 to 20 iterations when a suitable initial guess (see [50]) has been provided. However, when the gap between the highest occupied and lowest unoccupied orbital is small, which commonly occurs in, e.g., first-row transition metal complexes, the SCF procedure may become extremely slow to converge, oscillate between two or more solutions, converge to a higher-lying solution, or even to a saddle-point solution. Namely, it is critically important to realize that even if the orbital gradient vanishes, or equivalently, that the SCF equations are fulfilled, this does not mean that the energy expression Equation (7) truly has been minimized. Because there are typically several occupied as well as virtual orbitals, the minimization problem involves a large number of degrees of freedom. In multivariate calculus, a vanishing gradient only means that the orbitals correspond to some kind of extremum of the energy: a local minimum, a saddle-point solution, or even a local maximum, although the lattermost is highly improbable in SCF calculations.

In contrast to the sometimes erratic behavior of the SCF method, direct minimization based on orbital rotations is guaranteed to converge onto an extremal point $f_{ia} = 0$ per the theory of numerical analysis; this is of great worth when studying systems with complicated electronic structures for which conventional SCF algorithms fail. However, more predictable convergence does not come for free: the downside of direct minimization methods is that they carry a higher computational cost due to, e.g., the use of line searches in the orbital optimization. Direct minimization methods can also be formulated at the second order, yielding more robust convergence to a local minimum at the cost of more computational resources per iteration, see, e.g., [61–63]. Because direct minimization methods are based on an explicit rotation of the orbitals, they are able to always follow the same solution at variance to SCF methods where the orbital occupations are typically reset at every iteration according to the Aufbau principle. Because of this, direct minimization can lead to a solution where the Aufbau rule is violated, that is, the highest occupied orbital lies higher in energy than the lowest unoccupied orbital. Direct minimization methods can also be straightforwardly applied in more complicated electronic structure theories than self-consistent field theory. Such methods

may especially include explicit dependence on the molecular orbitals, as discussed by one of the present authors in [31,64,65] for the Perdew–Zunger self-interaction correction [66] which depends explicitly on the N_σ occupied orbitals, and [67] for the perfect quadruples [68] and perfect hexuples [69] models that also depend on the N_σ corresponding virtual orbitals.

In order to check the character of the extremum found by the SCF procedure or a direct minimization method, it is necessary to continue the analysis to second-order changes in the energy with respect to the orbital rotations by finding the lowest eigenvalue of the Hessian matrix: if it is negative, rotating the orbitals in the direction of the corresponding eigenvector will result in a further decrease of the energy. Whenever post-HF calculations are performed, or benchmark-quality values are sought at the SCF level, stability analysis [70,71] should be used to guarantee that the wave function indeed corresponds to a local minimum. Alternatively, trust-region methods [72–74] can be employed to ensure that the orbitals converge onto a true local minimum.

As always in the minimization of multivariate functions, locating the global minimum is difficult, and typically the best one can hope for is to find a local minimum. Some systems permit several local electronic minima: for instance, charge transfer complexes may allow both a neutral ($X \cdots Y$) as well as an ionic ($X^+ \cdots Y^-$) solution. Finding such physically motivated solutions is often straightforward by suitable manipulations of the initial guess, for instance, by constructing guesses via the superposition of atomic potentials [75,76] with the correct atomic charges. Sometimes it may also be interesting to locate saddle-point solutions, which have physical interpretations as excited states. Specific excited states can be explored within the SCF approach by replacing Aufbau population of the orbitals with overlap criteria [77,78] or with direct minimization by replacing the energy with the square of the gradient [79]; for instance, such an approach has been recently shown to predict highly accurate core spectra [80]. The full space of SCF solutions can be explored via, e.g., meta-dynamics [81].

11. Density Functional Contributions to Kohn–Sham–Fock Matrix

In Section 6 we derived expressions for the Kohn–Sham–Fock matrix elements for all but the density functional contribution

$$E^{XC} = b \int f(\vec{r}) d\vec{r} \quad (47)$$

which we will consider next. Hundreds of density functionals $f(\vec{r})$ of various forms have been published in the literature in the recent decades [82], and offering a comprehensive selection thereof poses a considerable challenge to quantum chemistry software developers. This problem is further exacerbated by the need to keep track with the several new functionals still being published every year. Moreover, the density functionals $f(\vec{r})$ typically carry extremely complicated functional forms, making their correct implementation painstaking work. The implementation is made even more difficult by the need to compute the first derivatives of $f(\vec{r})$ for the SCF procedure, as well as several higher-order ones for, e.g., the calculation of various properties.

Fortunately, these challenges have been obviated by freely available, portable standard implementations such as LIBXC [83] and XCFUN [84]. The LIBXC software package strives to implement *all DFT functionals published in the literature*, and provides a uniform interface to ~ 500 functionals of various forms. At present, LIBXC is used by ~ 30 electronic structure programs based on various numerical representations that range from basis set approaches (Gaussian-type orbitals, Slater-type orbitals, numerical atomic orbitals, finite elements, plane waves) to finite difference procedures. New functionals only have to be added once to LIBXC, meaning the library is easily kept up to date, after which they become available to all programs that support the corresponding rung of functionals on Jacob’s ladder [28]. Next, we will derive the equations necessary to implement the various rungs’ functionals in the variational basis set approach.

11.1. LDA Functionals

The simplest density functional approximations (DFAs), belonging to the first rung of Jacob's Ladder [28], are generally referred to as local (spin) density approximations (LDAs). These are functions of only the electron density [13]

$$f(\vec{r}) = f(\rho_\alpha(\vec{r}), \rho_\beta(\vec{r})) \quad (48)$$

such as the LDA exchange functional [85,86]

$$f(\vec{r}) = -\frac{3}{2} \left(\frac{3}{4\pi} \right)^{\frac{1}{3}} \left(\rho_\alpha^{4/3}(\vec{r}) + \rho_\beta^{4/3}(\vec{r}) \right)$$

Assuming f has the form of Equation (48), the resulting contribution to the Kohn–Sham–Fock matrix $F_{\mu\nu}^{\alpha;\text{XC}} = \partial E^{\text{XC}} / \partial P_{\mu\nu}^\alpha$ can be evaluated using Equation (5) for the densities at point \vec{r}

$$\rho_\sigma(\vec{r}) = \sum_{\mu\nu} P_{\mu\nu}^\sigma \chi_\mu(\vec{r}) \chi_\nu(\vec{r}) \quad (49)$$

as [16]

$$F_{\mu\nu}^{\alpha;\text{LDA}} = \frac{\partial E^{\text{XC}}}{\partial P_{\mu\nu}^\alpha} = b \int d\vec{r} \frac{\partial f(\vec{r})}{\partial \rho_\alpha} \frac{\partial \rho_\alpha}{\partial P_{\mu\nu}^\alpha} = b \int d\vec{r} \frac{\partial f(\vec{r})}{\partial \rho_\alpha} \chi_\mu(\vec{r}) \chi_\nu(\vec{r}) \quad (50)$$

with $\mathbf{F}^{\beta;\text{LDA}}$ having an analogous expression. Note that if the integral is evaluated using numerical quadrature,

$$F_{\mu\nu}^{\alpha;\text{LDA}} \approx b \sum_i w_i \frac{\partial f(\vec{r}_i)}{\partial \rho_\alpha} \chi_\mu(\vec{r}_i) \chi_\nu(\vec{r}_i) \quad (51)$$

Becke's multigrid approach [87] and further developments thereof being the standard approach in LCAO programs, the expression of Equation (51) can be most efficiently formulated with matrix products. Storing the values of the basis functions at the quadrature points as a matrix

$$X_{\mu i} = \chi_\mu(\vec{r}_i)$$

and defining a scaled version thereof as

$$Y_{\mu i}^\alpha = w_i \frac{\partial f(\vec{r}_i)}{\partial \rho_\alpha} \chi_\mu(\vec{r}_i)$$

the Fock matrix contribution can be evaluated as simply as

$$\mathbf{F}^{\alpha;\text{LDA}} = b \mathbf{X} (\mathbf{Y}^\alpha)^T$$

which is orders of magnitude faster than a simple *for* loop based algorithm.

11.2. GGA Functionals

The second rung of Jacob's Ladder [28] is referred to as the Generalised Gradient Approximation [88] (GGA). Density functional approximations on this rung also depend on the derivatives of the density

$$f(\vec{r}) = f(\rho_\alpha, \rho_\beta, \gamma_{\alpha\alpha}, \gamma_{\alpha\beta}, \gamma_{\beta\beta})$$

via the reduced gradients $\gamma_{\alpha\alpha} = \nabla\rho_\alpha \cdot \nabla\rho_\alpha$, $\gamma_{\alpha\beta} = \nabla\rho_\alpha \cdot \nabla\rho_\beta$, and $\gamma_{\beta\beta} = \nabla\rho_\beta \cdot \nabla\rho_\beta$, with the gradient of the density $\nabla\rho_\sigma$ being determined by

$$\nabla\rho_\sigma = \sum_{\mu\nu} P_{\mu\nu}^\sigma \nabla[\chi_\mu(\vec{r})\chi_\nu(\vec{r})] = \sum_{\mu\nu} P_{\mu\nu}^\sigma [\chi_\nu(\vec{r})\nabla\chi_\mu(\vec{r}) + \chi_\mu(\vec{r})\nabla\chi_\nu(\vec{r})]$$

The GGA contribution to the Fock matrix is given by [16]

$$\begin{aligned} F_{\mu\nu}^{\alpha;\text{GGA}} &= \frac{\partial E^{\text{XC}}}{\partial P_{\mu\nu}^\alpha} \\ &= \int d\vec{r} \left[\frac{\partial f(\vec{r})}{\partial \rho_\alpha} \frac{\partial \rho_\alpha}{\partial P_{\mu\nu}^\alpha} + \frac{\partial f(\vec{r})}{\partial \gamma_{\alpha\alpha}} \frac{\partial \gamma_{\alpha\alpha}}{\partial \nabla\rho_\alpha} \cdot \frac{\partial \nabla\rho_\alpha}{\partial P_{\mu\nu}^\alpha} + \frac{\partial f(\vec{r})}{\partial \gamma_{\alpha\beta}} \frac{\partial \gamma_{\alpha\beta}}{\partial \nabla\rho_\alpha} \cdot \frac{\partial \nabla\rho_\alpha}{\partial P_{\mu\nu}^\alpha} \right] = F_{\mu\nu}^{\alpha;\text{LDA}} \\ &+ \int d\vec{r} \left(2 \frac{\partial f(\vec{r})}{\partial \gamma_{\alpha\alpha}} \nabla\rho_\alpha(\vec{r}) + \frac{\partial f(\vec{r})}{\partial \gamma_{\alpha\beta}} \nabla\rho_\beta(\vec{r}) \right) \cdot [\chi_\nu(\vec{r})\nabla\chi_\mu(\vec{r}) + \chi_\mu(\vec{r})\nabla\chi_\nu(\vec{r})] \end{aligned} \quad (52)$$

The β expression can be obtained by switching α and β in Equation (52). In the restricted case, $E^{\text{XC}} = \int d\vec{r} f(\rho, \gamma)$ with $\gamma = \nabla\rho \cdot \nabla\rho$, which leads to the DFT contributions to F^{GGA} given by the simpler expression

$$F_{\mu\nu}^{\text{GGA}} = \int d\vec{r} \left[\frac{\partial f(\vec{r})}{\partial \rho} \chi_\mu(\vec{r})\chi_\nu(\vec{r}) + 2 \frac{\partial f(\vec{r})}{\partial \gamma} \nabla\rho(\vec{r}) \cdot [\chi_\nu(\vec{r})\nabla\chi_\mu(\vec{r}) + \chi_\mu(\vec{r})\nabla\chi_\nu(\vec{r})] \right] \quad (53)$$

Practical implementations of Equations (52) and (53) can again be formulated using matrix products.

11.3. Meta-GGA Functionals

On the third rung on Jacob's Ladder [28] are the meta-GGA (mGGA) approximations

$$f(\vec{r}) = f(\rho_\alpha, \rho_\beta, \gamma_{\alpha\alpha}, \gamma_{\alpha\beta}, \gamma_{\beta\beta}, \tau_\alpha, \tau_\beta, \nabla^2\rho_\alpha, \nabla^2\rho_\beta)$$

in which τ_σ and $\nabla^2\rho_\sigma$ are obtained as

$$\begin{aligned} \tau_\sigma &= \frac{1}{2} \sum_{i=1}^{N_\sigma} |\nabla\varphi_i(\vec{r})|^2 = \frac{1}{2} \sum_{i=1}^{N_\sigma} \nabla\varphi_i(\vec{r}) \cdot \nabla\varphi_i(\vec{r}) = \frac{1}{2} \sum_{i=1}^{N_\sigma} \sum_{\mu\nu} C_{\mu i}^\sigma C_{\nu i}^\sigma \nabla\chi_\mu(\vec{r}) \cdot \nabla\chi_\nu(\vec{r}) \\ &= \frac{1}{2} \sum_{\mu\nu} P_{\mu\nu}^\sigma \nabla\chi_\mu(\vec{r}) \cdot \nabla\chi_\nu(\vec{r}) \end{aligned} \quad (54)$$

$$\begin{aligned} \nabla^2\rho_\sigma &= \sum_{\mu\nu} P_{\mu\nu}^\sigma \nabla^2[\chi_\mu(\vec{r})\chi_\nu(\vec{r})] \\ &= \sum_{\mu\nu} P_{\mu\nu}^\sigma [\chi_\mu(\vec{r})\nabla^2\chi_\nu(\vec{r}) + 2\nabla\chi_\mu(\vec{r}) \cdot \nabla\chi_\nu(\vec{r}) + \chi_\nu(\vec{r})\nabla^2\chi_\mu(\vec{r})] \\ &= \sum_{\mu\nu} P_{\mu\nu}^\sigma [\chi_\mu(\vec{r})\nabla^2\chi_\nu(\vec{r}) + \chi_\nu(\vec{r})\nabla^2\chi_\mu(\vec{r})] + 4\tau_\sigma \end{aligned} \quad (55)$$

The meta-GGA contributions to the Kohn–Sham–Fock matrix are straightforwardly obtained as [20]

$$\begin{aligned}
 F_{\mu\nu}^{\alpha;\text{mGGA}} &= \frac{\partial E^{\text{XC}}}{\partial P_{\mu\nu}^{\alpha}} \\
 &= F_{\mu\nu}^{\alpha;\text{GGA}} + \int d\vec{r} \left[\frac{\partial f(\vec{r})}{\partial \tau_{\alpha}} \frac{\partial \tau_{\alpha}}{\partial P_{\mu\nu}^{\alpha}} + \frac{\partial f(\vec{r})}{\partial \nabla^2 \rho_{\alpha}(\vec{r})} \frac{\partial \nabla^2 \rho_{\alpha}(\vec{r})}{\partial P_{\mu\nu}^{\alpha}} \right] \\
 &= F_{\mu\nu}^{\alpha;\text{GGA}} + \int d\vec{r} \left[\left(\frac{1}{2} \frac{\partial f(\vec{r})}{\partial \tau_{\alpha}} + 2 \frac{\partial f(\vec{r})}{\partial \nabla^2 \rho_{\alpha}(\vec{r})} \right) \nabla \chi_{\mu}(\vec{r}) \cdot \nabla \chi_{\nu}(\vec{r}) \right. \\
 &\quad \left. + \frac{\partial f(\vec{r})}{\partial \nabla^2 \rho_{\alpha}(\vec{r})} [\chi_{\mu}(\vec{r}) \nabla^2 \chi_{\nu}(\vec{r}) + \chi_{\nu}(\vec{r}) \nabla^2 \chi_{\mu}(\vec{r})] \right]
 \end{aligned} \tag{56}$$

which can again be expressed in terms of matrix products to achieve faster quadrature. The expressions remain formally the same in the restricted case, but the quantities correspond to the total electron density.

11.4. Range-Separated Hybrid Functionals

As was mentioned before, the use of non-zero values for the constants a and b in Equation (7) allows the inclusion of exact exchange effects in a DFT calculation. Such functionals represent the fourth rung of Jacob’s Ladder [28], and are generally referred to as hybrids. A further development on hybrid functionals are range-separated hybrids [89,90], in which the interelectronic interaction is divided into a short-range (sr) and a long-range (lr) part with a resolution of the identity

$$\frac{1}{r_{12}} = \frac{\phi^{\text{sr}}(r_{12})}{r_{12}} + \frac{\phi^{\text{lr}}(r_{12})}{r_{12}} \tag{57}$$

where $\phi^{\text{sr}}(r_{12}) + \phi^{\text{lr}}(r_{12}) = 1$. The rationale for range separation is that since density functional approximations for the exchange are based only on local information about the density, they fail to reproduce accurate estimates for, e.g., charge transfer processes. Separating the interaction by range per Equation (57) leads to a hybrid exchange functional that has four contributions

$$E^{\text{X}} = a^{\text{sr}} E^{\text{sr-HF}} + a^{\text{lr}} E^{\text{lr-HF}} + b^{\text{sr}} E^{\text{sr-DFT}} + b^{\text{lr}} E^{\text{lr-DFT}} = a^{\text{sr}} E^{\text{sr-HF}} + a^{\text{lr}} E^{\text{lr-HF}} + E^{\text{DFT}} \tag{58}$$

where we have stressed that since the DFT contributions are evaluated based only on the density (and possibly its derivatives), $b^{\text{sr}} E^{\text{sr-DFT}} + b^{\text{lr}} E^{\text{lr-DFT}}$ is nothing but a definition of a new density functional. In contrast, the HF contributions to the energy and the Kohn–Sham–Fock matrix have to be evaluated separately with range-separated ERIs

$$(\mu\nu|\lambda\sigma)^{\text{sr/lr}} = \int d\vec{r}_1 \int d\vec{r}_2 \chi_{\mu}(\vec{r}_1) \chi_{\nu}(\vec{r}_1) \frac{\phi^{\text{sr/lr}}(r_{12})}{r_{12}} \chi_{\lambda}(\vec{r}_2) \chi_{\sigma}(\vec{r}_2) \tag{59}$$

Several kinds of range-separation kernels $\phi^{\text{sr/lr}}(r)$ have been proposed; however, the error function based kernel $\phi^{\text{sr}}(r;\omega) = \text{erfc}(\omega r)$, $\phi^{\text{lr}}(r;\omega) = \text{erf}(\omega r)$, where ω is the range-separation parameter, is by far the most commonly used one because it is exceedingly simple to implement in codes employing a plane-wave or Gaussian basis set [91,92]. The error function kernel is used, for instance, in the Heyd–Scuseria–Ernzerhof (HSE) functionals for solid-state calculations [91,93], as well as in the ω B97M-V [94] functional that is discussed below in Section 11.5. Some functionals based on Yukawa kernels, $\phi^{\text{sr}}(r;\omega) = e^{-\lambda r}$, $\phi^{\text{lr}}(r;\omega) = 1 - e^{-\lambda r}$, have also been published and are available in LIBXC, for instance. It is important to check that the range-separation kernel used in the density functional

implementation matches the one used in the computation of the range-separated ERIs in Equation (59), as the results will be incorrect otherwise.

11.5. Non-Local Correlation

Dispersion effects, i.e., van der Waals interactions, can be modeled in an *ab initio* DFT setting with non-local correlation functionals [95]

$$E^{\text{nlc}} = \int d\vec{r}_1 d\vec{r}_2 \rho(\vec{r}_1) \Phi_0(\vec{r}_1, \vec{r}_2) \rho(\vec{r}_2) \quad (60)$$

Because the non-local correlation energy term depends explicitly on the electron density, it also needs to be included in the SCF procedure, in principle. In contrast, empirical dispersion corrections such as Grimme's various DFT-D approaches [96–98] do not depend on the electron density, and are added only as an *ad hoc* correction onto the electronic energy.

Perhaps the most accurate rung-3 and rung-4 functionals currently available [99–101], the pure B97M-V [102] mGGA as well as the range-separated ω B97M-V [94] hybrid mGGA, respectively, are built on top of [103] the VV10 non-local correlation functional [23] which is controlled by two adjustable parameters, b and C , which have been trained alongside the density functional in B97M-V and ω B97M-V. The results of a recent benchmark study suggest that the VV10 contributions on densities and orbital energies are negligible, and that sufficiently accurate energetics may be obtained by a one-shot evaluation of E^{nlc} in a post-SCF fashion [99]. Still, a rigorous minimization of the energy requires considering the effects of the non-local correlation on the wave function. Although Equation (60) does not appear to fit on the rungs of functionals discussed above, the VV10 kernel turns out to yield a GGA-type contribution to the Kohn–Sham–Fock matrix as discussed in [23], to which we refer for further details.

12. Summary and Discussion

We have presented an overview of self-consistent field calculations within a variational basis set formalism, and discussed the solution of the self-consistent field equations arising from Hartree–Fock as well as various levels of density functional approximations using either the traditional fixed-point equations or direct minimization, as well as various conceptual and numerical issues arising in their implementation. No assumptions have been made on the underlying basis set in the present work: the self-consistent field formalism is the same regardless of the form of the basis functions, which can be chosen to be, e.g., Gaussian-type orbitals (GTOs), Slater-type orbitals (STOs), numerical atomic orbitals (NAOs), or finite element shape functions. The basis set is only reflected in the molecular integrals, that is, the matrix elements of the core Hamiltonian and the two-electron integrals. The various choices for the basis set have different advantages and disadvantages, including the evaluation of the molecular integrals; see [2] for discussion. Instead, the main restriction in the present work is the implicit assumption that the basis set is compact enough so that the $N \times N$ density and Fock matrices are small enough to allow $O(N^2)$ dense matrix storage and $O(N^3)$ diagonalization. Although the present overview has focused on non-relativistic calculations on molecules, the discussion for relativistic calculations as well as crystalline systems is analogous to a large degree. We hope that the present, consistent and thorough derivation and analysis will be useful for reference as well as teaching purposes, and that the results presented herein will lead to a wider availability of density functionals in electronic structure programs.

Author Contributions: S.L. derived the equations in section 11. C.V.A. and F.B. derived the equations in Sections 2–6 and wrote an initial version of the manuscript. S.L. wrote the final version of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work has been supported by the Academy of Finland (Suomen Akatemia) through project number 311149.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Poree, C.; Schoenebeck, F. A Holy Grail in Chemistry: Computational Catalyst Design: Feasible or Fiction? *Acc. Chem. Res.* **2017**, *50*, 605–608. [[CrossRef](#)]
2. Lehtola, S. A review on non-relativistic, fully numerical electronic structure calculations on atoms and diatomic molecules. *Int. J. Quantum Chem.* **2019**, *119*, e25968. [[CrossRef](#)]
3. Lehtola, S. Fully numerical Hartree–Fock and density functional calculations. I. Atoms. *Int. J. Quantum Chem.* **2019**, *119*, e25945. [[CrossRef](#)]
4. Lehtola, S. Fully numerical Hartree–Fock and density functional calculations. II. Diatomic molecules. *Int. J. Quantum Chem.* **2019**, *119*, e25944. [[CrossRef](#)]
5. Roothaan, C. New Developments in Molecular Orbital Theory. *Rev. Mod. Phys.* **1951**, *23*, 69–89. [[CrossRef](#)]
6. Hall, G.G. The Molecular Orbital Theory of Chemical Valency. VIII. A Method of Calculating Ionization Potentials. *Proc. R. Soc. A Math. Phys. Eng. Sci.* **1951**, *205*, 541–552. [[CrossRef](#)]
7. Löwdin, P.O. On the Nonorthogonality Problem. *Adv. Quantum Chem.* **1970**, *5*, 185–199. [[CrossRef](#)]
8. Pople, J.A.; Nesbet, R.K. Self-Consistent Orbitals for Radicals. *J. Chem. Phys.* **1954**, *22*, 571. [[CrossRef](#)]
9. Berthier, G. Configurations électroniques incomplètes. Partie I. La Méthode du Champ Moléculaire Self-Consistent et l’Etude des Etats à Couches Incomplètes. *J. Chim. Phys.* **1954**, *51*, 363–371. [[CrossRef](#)]
10. Roothaan, C. Self-Consistent Field Theory for Open Shells of Electronic Systems. *Rev. Mod. Phys.* **1960**, *32*, 179–185. [[CrossRef](#)]
11. Krebs, S. A review on the derivation of the spin-Restricted Hartree–Fock (RHF) Self-Consistent Field (SCF) equations for open-shell systems. Description of different methods to handle the off-diagonal Lagrangian multipliers coupling closed and open shells. *Comput. Phys. Commun.* **1999**, *116*, 137–277. [[CrossRef](#)]
12. Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136*, B864–B871. [[CrossRef](#)]
13. Kohn, W.; Sham, L.J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140*, A1133–A1138. [[CrossRef](#)]
14. Baerends, E.J. Perspective on “Self-consistent equations including exchange and correlation effects”. In *Theoretical Chemistry Accounts*; Cramer, C.J., Truhlar, D.G., Eds.; Springer: Berlin/Heidelberg, Germany, 2000; pp. 265–269. [[CrossRef](#)]
15. Geerlings, P.; De Proft, F.; Langenaeker, W. Conceptual Density Functional Theory. *Chem. Rev.* **2003**, *103*, 1793–1874. [[CrossRef](#)] [[PubMed](#)]
16. Pople, J.A.; Gill, P.M.W.; Johnson, B.G. Kohn–Sham density-functional theory within a finite basis set. *Chem. Phys. Lett.* **1992**, *199*, 557–560. [[CrossRef](#)]
17. Johnson, B.G.; Gill, P.M.W.; Pople, J.A. The performance of a family of density functional methods. *J. Chem. Phys.* **1993**, *98*, 5612. [[CrossRef](#)]
18. Johnson, B.G. Erratum: The performance of a family of density functional methods. *J. Chem. Phys.* **1994**, *101*, 9202. [[CrossRef](#)]
19. Kobayashi, K.; Kurita, N.; Kumahora, H.; Tago, K. Bond-energy calculations of Cu₂, Ag₂, and CuAg with the generalized gradient approximation. *Phys. Rev. A* **1991**, *43*, 5810–5813. [[CrossRef](#)]
20. Neumann, R.; Nobes, R.H.; Handy, N.C. Exchange functionals and potentials. *Mol. Phys.* **1996**, *87*, 1–36. [[CrossRef](#)]
21. Vydrov, O.A.; Wu, Q.; Van Voorhis, T. Self-consistent implementation of a nonlocal van der Waals density functional with a Gaussian basis set. *J. Chem. Phys.* **2008**, *129*. [[CrossRef](#)]
22. Vydrov, O.A.; Van Voorhis, T. Improving the accuracy of the nonlocal van der Waals density functional with minimal empiricism. *J. Chem. Phys.* **2009**, *130*, 104105. [[CrossRef](#)] [[PubMed](#)]

23. Vydrov, O.A.; Van Voorhis, T. Nonlocal van der Waals density functional: The simpler the better. *J. Chem. Phys.* **2010**, *133*, 244103. [[CrossRef](#)] [[PubMed](#)]
24. Vydrov, O.A.; Van Voorhis, T. Implementation and assessment of a simple nonlocal van der Waals density functional. *J. Chem. Phys.* **2010**, *132*. [[CrossRef](#)] [[PubMed](#)]
25. Parrish, R.M.; Burns, L.A.; Smith, D.G.A.; Simmonett, A.C.; DePrince, A.E.; Hohenstein, E.G.; Bozkaya, U.; Sokolov, A.Y.; Di Remigio, R.; Richard, R.M.; et al. Psi4 1.1: An Open-Source Electronic Structure Program Emphasizing Automation, Advanced Libraries, and Interoperability. *J. Chem. Theory Comput.* **2017**, *13*, 3185–3197. [[CrossRef](#)] [[PubMed](#)]
26. Sun, Q.; Berkelbach, T.C.; Blunt, N.S.; Booth, G.H.; Guo, S.; Li, Z.; Liu, J.; McClain, J.D.; Sayfutyarova, E.R.; Sharma, S.; et al. PySCF: The Python-based simulations of chemistry framework. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2018**, *8*, e1340. [[CrossRef](#)]
27. Becke, A.D.; Roussel, M.R. Exchange holes in inhomogeneous systems: A coordinate-space model. *Phys. Rev. A* **1989**, *39*, 3761–3767. [[CrossRef](#)]
28. Perdew, J.P. Jacob's ladder of density functional approximations for the exchange-correlation energy. *AIP Conf. Proc.* **2001**, *577*, 1–20. [[CrossRef](#)]
29. Small, D.W.; Sundstrom, E.J.; Head-Gordon, M. Restricted Hartree Fock using complex-valued orbitals: A long-known but neglected tool in electronic structure theory. *J. Chem. Phys.* **2015**, *142*, 024104. [[CrossRef](#)]
30. Lee, J.; Bertels, L.W.; Small, D.W.; Head-Gordon, M. Kohn–Sham Density Functional Theory with Complex, Spin-Restricted Orbitals: Accessing a New Class of Densities without the Symmetry Dilemma. *Phys. Rev. Lett.* **2019**, *123*, 113001. [[CrossRef](#)]
31. Lehtola, S.; Head-Gordon, M.; Jónsson, H. Complex Orbitals, Multiple Local Minima, and Symmetry Breaking in Perdew–Zunger Self-Interaction Corrected Density Functional Theory Calculations. *J. Chem. Theory Comput.* **2016**, *12*, 3195–3207. [[CrossRef](#)]
32. Perdew, J.P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868. [[CrossRef](#)] [[PubMed](#)]
33. Becke, A.D. A new mixing of Hartree–Fock and local density-functional theories. *J. Chem. Phys.* **1993**, *98*, 1372–1377. [[CrossRef](#)]
34. Stephens, P.J.; Devlin, F.J.; Chabalowski, C.F.; Frisch, M.J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem.* **1994**, *98*, 11623–11627. [[CrossRef](#)]
35. Lehtola, S.; Jónsson, H. Unitary Optimization of Localized Molecular Orbitals. *J. Chem. Theory Comput.* **2013**, *9*, 5365–5372. [[CrossRef](#)]
36. Foster, J.; Boys, S. Canonical Configurational Interaction Procedure. *Rev. Mod. Phys.* **1960**, *32*, 300–302. [[CrossRef](#)]
37. Edmiston, C.; Ruedenberg, K. Localized Atomic and Molecular Orbitals. *Rev. Mod. Phys.* **1963**, *35*, 457–464. [[CrossRef](#)]
38. Pipek, J.; Mezey, P.G. A fast intrinsic localization procedure applicable for ab initio and semiempirical linear combination of atomic orbital wave functions. *J. Chem. Phys.* **1989**, *90*, 4916. [[CrossRef](#)]
39. Høyvik, I.; Jansik, B.; Jørgensen, P. Orbital localization using fourth central moment minimization. *J. Chem. Phys.* **2012**, *137*, 224114. [[CrossRef](#)]
40. Lehtola, S.; Jónsson, H. Pipek–Mezey Orbital Localization Using Various Partial Charge Estimates. *J. Chem. Theory Comput.* **2014**, *10*, 642–649. [[CrossRef](#)]
41. Knizia, G.; Klein, J.E.M.N. Electron Flow in Reaction Mechanisms-Revealed from First Principles. *Angew. Chemie Int. Ed.* **2015**, *54*, 5518–5522. [[CrossRef](#)]
42. Liu, Y.; Kilby, P.; Frankcombe, T.J.; Schmidt, T.W. Calculating curly arrows from ab initio wavefunctions. *Nat. Commun.* **2018**, *9*, 1436. [[CrossRef](#)] [[PubMed](#)]
43. Klein, J.E.M.N.; Knizia, G.; Rzepa, H.S. Epoxidation of Alkenes by Peracids: From Textbook Mechanisms to a Quantum Mechanically Derived Curly-Arrow Depiction. *ChemistryOpen* **2019**, *8*, 1244–1250. [[CrossRef](#)] [[PubMed](#)]

44. Head-Gordon, M.; Pople, J.A. Optimization of wave function and geometry in the finite basis Hartree–Fock method. *J. Phys. Chem.* **1988**, *92*, 3063–3069. [[CrossRef](#)]
45. Van Voorhis, T.; Head-Gordon, M. A geometric approach to direct minimization. *Mol. Phys.* **2002**, *100*, 1713–1721. [[CrossRef](#)]
46. Coulson, C.A.; Fischer, I. XXXIV. Notes on the molecular orbital treatment of the hydrogen molecule. *Lond. Edinb. Dublin Philos. Mag. J. Sci.* **1949**, *40*, 386–393. [[CrossRef](#)]
47. Löwdin, P.O. Quantum theory of cohesive properties of solids. *Adv. Phys.* **1956**, *5*, 1–171. [[CrossRef](#)]
48. Lehtola, S. Curing basis set overcompleteness with pivoted Cholesky decompositions. *J. Chem. Phys.* **2019**, *151*, 241102. [[CrossRef](#)]
49. Lehtola, S. Accurate reproduction of strongly repulsive interatomic potentials. *Phys. Rev. A.* **2020**, *101*, 032504. [[CrossRef](#)]
50. Lehtola, S. Assessment of Initial Guesses for Self-Consistent Field Calculations. Superposition of Atomic Potentials: Simple yet Efficient. *J. Chem. Theory Comput.* **2019**, *15*, 1593–1604. [[CrossRef](#)]
51. Karlström, G. Dynamical damping based on energy minimization for use ab initio molecular orbital SCF calculations. *Chem. Phys. Lett.* **1979**, *67*, 348–350. [[CrossRef](#)]
52. Cancès, E.; Le Bris, C. Can we outperform the DIIS approach for electronic structure calculations? *Int. J. Quantum Chem.* **2000**, *79*, 82–90. [[CrossRef](#)]
53. Saunders, V.R.; Hillier, I.H. A “Level-Shifting” method for converging closed shell Hartree–Fock wave functions. *Int. J. Quantum Chem.* **1973**, *7*, 699–705. [[CrossRef](#)]
54. Mitin, A.V. The dynamic “level shift” method for improving the convergence of the SCF procedure. *J. Comput. Chem.* **1988**, *9*, 107–110. [[CrossRef](#)]
55. Dömötör, G.; Bán, M.I. Dynamic level-shifting. *Comput. Chem.* **1989**, *13*, 53–57. [[CrossRef](#)]
56. Pulay, P. Convergence acceleration of iterative sequences. The case of SCF iteration. *Chem. Phys. Lett.* **1980**, *73*, 393–398. [[CrossRef](#)]
57. Pulay, P. Improved SCF convergence acceleration. *J. Comput. Chem.* **1982**, *3*, 556–560. [[CrossRef](#)]
58. Kudin, K.N.; Scuseria, G.E.; Cancès, E. A black-box self-consistent field convergence algorithm: One step closer. *J. Chem. Phys.* **2002**, *116*, 8255. [[CrossRef](#)]
59. Hu, X.; Yang, W. Accelerating self-consistent field convergence with the augmented Roothaan–Hall energy function. *J. Chem. Phys.* **2010**, *132*, 054109. [[CrossRef](#)]
60. Rabuck, A.D.; Scuseria, G.E. Improving self-consistent field convergence by varying occupation numbers. *J. Chem. Phys.* **1999**, *110*, 695–700. [[CrossRef](#)]
61. Douady, J.; Ellinger, Y.; Subra, R.; Levy, B. A quadratically convergent SCF procedure. *Comput. Phys. Commun.* **1979**, *17*, 23–25. [[CrossRef](#)]
62. Douady, J.; Ellinger, Y.; Subra, R.; Levy, B. Exponential transformation of molecular orbitals: A quadratically convergent SCF procedure. I. General formulation and application to closed-shell ground states. *J. Chem. Phys.* **1980**, *72*, 1452. [[CrossRef](#)]
63. Head-Gordon, M.; Pople, J.A.; Frisch, M.J. Quadratically convergent simultaneous optimization of wavefunction and geometry. *Int. J. Quantum Chem.* **1989**, *36*, 291–303. [[CrossRef](#)]
64. Lehtola, S.; Jónsson, H. Variational, Self-Consistent Implementation of the Perdew–Zunger Self-Interaction Correction with Complex Optimal Orbitals. *J. Chem. Theory Comput.* **2014**, *10*, 5324–5337. [[CrossRef](#)] [[PubMed](#)]
65. Lehtola, S.; Jónsson, H. Correction to Variational, Self-Consistent Implementation of the Perdew–Zunger Self-Interaction Correction with Complex Optimal Orbitals. *J. Chem. Theory Comput.* **2015**, *11*, 5052–5053. [[CrossRef](#)]
66. Perdew, J.P.; Zunger, A. Self-interaction correction to density-functional approximations for many-electron systems. *Phys. Rev. B* **1981**, *23*, 5048–5079. [[CrossRef](#)]
67. Lehtola, S.; Parkhill, J.; Head-Gordon, M. Orbital optimisation in the perfect pairing hierarchy: Applications to full-valence calculations on linear polyacenes. *Mol. Phys.* **2018**, *116*, 547–560. [[CrossRef](#)]
68. Parkhill, J.A.; Lawler, K.; Head-Gordon, M. The perfect quadruples model for electron correlation in a valence active space. *J. Chem. Phys.* **2009**, *130*, 084101. [[CrossRef](#)]

69. Parkhill, J.A.; Head-Gordon, M. A tractable and accurate electronic structure method for static correlations: The perfect hexuples model. *J. Chem. Phys.* **2010**, *133*, 024103. [[CrossRef](#)]
70. Seeger, R.; Pople, J.A. Self-consistent molecular orbital methods. XVIII. Constraints and stability in Hartree–Fock theory. *J. Chem. Phys.* **1977**, *66*, 3045–3050. [[CrossRef](#)]
71. Bauernschmitt, R.; Ahlrichs, R. Stability analysis for solutions of the closed shell Kohn–Sham equation. *J. Chem. Phys.* **1996**, *104*, 9047. [[CrossRef](#)]
72. Francisco, J.B.; Martínez, J.M.; Martínez, L. Globally convergent trust-region methods for self-consistent field electronic structure calculations. *J. Chem. Phys.* **2004**, *121*, 10863–10878. [[CrossRef](#)] [[PubMed](#)]
73. Thøgersen, L.; Olsen, J.; Yeager, D.; Jørgensen, P.; Salek, P.; Helgaker, T. The trust-region self-consistent field method: Towards a black-box optimization in Hartree–Fock and Kohn–Sham theories. *J. Chem. Phys.* **2004**, *121*, 16–27. [[CrossRef](#)] [[PubMed](#)]
74. Thøgersen, L.; Olsen, J.; Köhn, A.; Jørgensen, P.; Salek, P.; Helgaker, T. The trust-region self-consistent field method in Kohn–Sham density-functional theory. *J. Chem. Phys.* **2005**, *123*, 074103. [[CrossRef](#)] [[PubMed](#)]
75. Almlöf, J.; Faegri, K.; Korsell, K. Principles for a direct SCF approach to LCAO-MO ab-initio calculations. *J. Comput. Chem.* **1982**, *3*, 385–399. [[CrossRef](#)]
76. Van Lenthe, J.H.; Zwaans, R.; Van Dam, H.J.J.; Guest, M.F. Starting SCF calculations by superposition of atomic densities. *J. Comput. Chem.* **2006**, *27*, 926–932. [[CrossRef](#)]
77. Gilbert, A.T.B.; Besley, N.A.; Gill, P.M.W. Self-Consistent Field Calculations of Excited States Using the Maximum Overlap Method (MOM). *J. Phys. Chem. A* **2008**, *112*, 13164–13171. [[CrossRef](#)]
78. Barca, G.M.J.; Gilbert, A.T.B.; Gill, P.M.W. Simple Models for Difficult Electronic Excitations. *J. Chem. Theory Comput.* **2018**, *14*, 1501–1509. [[CrossRef](#)]
79. Hait, D.; Head-Gordon, M. Excited state orbital optimization via minimizing the square of the gradient: General approach and application to singly and doubly excited states via density functional theory. *J. Chem. Theory Comput.* **2020**. [[CrossRef](#)]
80. Hait, D.; Head-Gordon, M. Highly Accurate Prediction of Core Spectra of Molecules at Density Functional Theory Cost: Attaining Sub-electronvolt Error from a Restricted Open-Shell Kohn–Sham Approach. *J. Phys. Chem. Lett.* **2020**, *11*, 775–786. [[CrossRef](#)]
81. Thom, A.; Head-Gordon, M. Locating Multiple Self-Consistent Field Solutions: An Approach Inspired by Metadynamics. *Phys. Rev. Lett.* **2008**, *101*, 193001. [[CrossRef](#)]
82. Mardirossian, N.; Head-Gordon, M. Thirty years of density functional theory in computational chemistry: An overview and extensive assessment of 200 density functionals. *Mol. Phys.* **2017**, *115*, 2315–2372. [[CrossRef](#)]
83. Lehtola, S.; Steigemann, C.; Oliveira, M.J.T.; Marques, M.A.L. Recent developments in LIBXC—A comprehensive library of functionals for density functional theory. *SoftwareX* **2018**, *7*, 1–5. [[CrossRef](#)]
84. Ekström, U.; Visscher, L.; Bast, R.; Thorvaldsen, A.J.; Ruud, K. Arbitrary-Order Density Functional Response Theory from Automatic Differentiation. *J. Chem. Theory Comput.* **2010**, *6*, 1971–1980. [[CrossRef](#)] [[PubMed](#)]
85. Bloch, F. Bemerkung zur Elektronentheorie des Ferromagnetismus und der elektrischen Leitfähigkeit. *Z. Phys.* **1929**, *57*, 545–555. [[CrossRef](#)]
86. Dirac, P.A.M. Note on Exchange Phenomena in the Thomas Atom. *Math. Proc. Cambridge Philos. Soc.* **1930**, *26*, 376–385. [[CrossRef](#)]
87. Becke, A.D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **1988**, *38*, 3098–3100. [[CrossRef](#)]
88. Perdew, J.P.; Yue, W. Accurate and simple density functional for the electronic exchange energy: Generalized gradient approximation. *Phys. Rev. B* **1986**, *33*, 8800–8802. [[CrossRef](#)]
89. Savin, A.; Flad, H.J. Density functionals for the Yukawa electron-electron interaction. *Int. J. Quantum Chem.* **1995**, *56*, 327–332. [[CrossRef](#)]
90. Leininger, T.; Stoll, H.; Werner, H.J.; Savin, A. Combining long-range configuration interaction with short-range density functionals. *Chem. Phys. Lett.* **1997**, *275*, 151–160. [[CrossRef](#)]
91. Heyd, J.; Scuseria, G.E.; Ernzerhof, M. Hybrid functionals based on a screened Coulomb potential. *J. Chem. Phys.* **2003**, *118*, 8207. [[CrossRef](#)]

92. Ahlrichs, R. A simple algebraic derivation of the Obara-Saika scheme for general two-electron interaction potentials. *Phys. Chem. Chem. Phys.* **2006**, *8*, 3072–3077. [[CrossRef](#)] [[PubMed](#)]
93. Heyd, J.; Scuseria, G.E.; Ernzerhof, M. Erratum: “Hybrid functionals based on a screened Coulomb potential”. *J. Chem. Phys.* **2006**, *124*, 219906. [[CrossRef](#)]
94. Mardirossian, N.; Head-Gordon, M. ω B97M-V: A combinatorially optimized, range-separated hybrid, meta-GGA density functional with VV10 nonlocal correlation. *J. Chem. Phys.* **2016**, *144*, 214110. [[CrossRef](#)] [[PubMed](#)]
95. Berland, K.; Cooper, V.R.; Lee, K.; Schröder, E.; Thonhauser, T.; Hyldgaard, P.; Lundqvist, B.I. van der Waals forces in density functional theory: A review of the vdW-DF method. *Rep. Prog. Phys.* **2015**, *78*, 066501. [[CrossRef](#)] [[PubMed](#)]
96. Grimme, S. Accurate description of van der Waals complexes by density functional theory including empirical corrections. *J. Comput. Chem.* **2004**, *25*, 1463–1473. [[CrossRef](#)] [[PubMed](#)]
97. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104. [[CrossRef](#)] [[PubMed](#)]
98. Caldeweyher, E.; Bannwarth, C.; Grimme, S. Extension of the D3 dispersion coefficient model. *J. Chem. Phys.* **2017**, *147*, 034112. [[CrossRef](#)]
99. Najibi, A.; Goerigk, L. The Nonlocal Kernel in van der Waals Density Functionals as an Additive Correction: An Extensive Analysis with Special Emphasis on the B97M-V and ω B97M-V Approaches. *J. Chem. Theory Comput.* **2018**, *14*, 5725–5738. [[CrossRef](#)]
100. Santra, G.; Sylvetsky, N.; Martin, J.M.L. Minimally Empirical Double-Hybrid Functionals Trained against the GMTKN55 Database: revDSD-PBEP86-D4, revDOD-PBE-D4, and DOD-SCAN-D4. *J. Phys. Chem. A* **2019**, *123*, 5129–5143. [[CrossRef](#)]
101. Iron, M.A.; Janes, T. Evaluating Transition Metal Barrier Heights with the Latest Density Functional Theory Exchange–Correlation Functionals: The MOBH35 Benchmark Database. *J. Phys. Chem. A* **2019**, *123*, 3761–3781. [[CrossRef](#)]
102. Mardirossian, N.; Head-Gordon, M. Mapping the genome of meta-generalized gradient approximation density functionals: The search for B97M-V. *J. Chem. Phys.* **2015**, *142*, 074111. [[CrossRef](#)] [[PubMed](#)]
103. Calbo, J.; Ortí, E.; Sancho-García, J.C.; Aragó, J. The Nonlocal Correlation Density Functional VV10. *Annu. Rep. Comput. Chem.* **2015**, *11*, 37–102. [[CrossRef](#)]

Sample Availability: Not available.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).