Density functional theory: Its origins, rise to prominence, and future

R. O. Jones*
Peter-Grünberg-Institut PGI-1 and German Research School for Simulation Sciences, Forschungszentrum Jülich, D-52425 Jülich, Germany
(published 25 August 2015)

In little more than 20 years, the number of applications of the density functional (DF) formalism in chemistry and materials science has grown in an astonishing fashion. The number of publications alone shows that DF calculations make up a huge success story, and many younger colleagues are surprised to learn that the widespread application of density functional methods, particularly in chemistry, began only after 1990. This is indeed unexpected, because the origins are usually traced to the papers of Hohenberg, Kohn, and Sham more than a quarter of a century earlier. The DF formalism, its applications, and prospects were reviewed for this journal in 1989. About the same time, the combination of DF calculations with molecular dynamics promised to provide an efficient way to study structures and reactions in molecules and extended systems. This paper reviews the development of density-related methods back to the early years of quantum mechanics and follows the breakthrough in their application after 1990. The two examples from biochemistry and materials science are among the many current applications that were simply far beyond expectations in 1990. The reasons why—50 years after its modern formulation and after two decades of rapid expansion—some of the most cited practitioners in the field are concerned about its future are discussed.

DOI: 10.1103/RevModPhys.87.897 PACS numbers: 71.15.Mb, 31.15.E—

CONTENTS

I. Introduction 897
II. The Density as a Basic Variable 898
III. An “Approximate Practical Method” 900
   A. Exchange hole and local density approximations
   B. Electron density and the chemical bond
IV. Modern Density Functional Formalism 902
   A. Single-particle description of a many-electron system
   B. Exchange-correlation energy $E_{xc}$
   C. Exchange-correlation hole and $E_{xc}$
V. Advances to 1990 905
   A. Theoretical progress
      1. Temperature, spin, and time dependence
      2. Developments of formalism
      3. Functional approximations
   B. Condensed matter
   C. Chemistry
   D. Menton, 1991
VI. After Acceptance (1990—Present) 910
   A. Progress and problems
      1. Approximations for $E_{xc}$
      2. Dispersion interactions
      3. Excitations and eigenvalues
      4. “Strongly correlated” systems
      5. Developments related to QMC
      6. Reduced density matrix functional theory
      7. “Missing” topics
      8. Terminology
   B. Two applications
C. DF theory in other contexts 915
   1. Classical density functional theory 915
   2. Nuclei 915

VII. Summary and Outlook 915
   A. An “approximate practical method” 915
   B. Exchange-correlation approximations 916
   C. Quo vadis? 916

Acknowledgments 917
Appendix: Adiabatic Coupling 917
References 917

I. INTRODUCTION

The density functional (DF) formalism shows that ground state (GS) and other properties of a system of electrons in an external field can be determined from knowledge of the electron density distribution $n(r)$ alone. Fermi (1927, 1928) and Thomas (1927) recognized the basic nature of the electron density and applied it to atoms, and Dirac (1930a) showed how exchange effects could be incorporated into this picture. In a far-sighted and largely overlooked observation, Dirac (1930a) also wrote that a “density function”—today we know it as the one-particle reduced density matrix—completely determines the whole state of the atom within the Hartree-Fock (HF) approximation; “it is not necessary to specify the individual three-dimensional wave functions.”

My main focus here is on a property for which DF calculations are particularly valuable in chemistry and materials science: the total energy $E$ of a system of electrons in the presence of ions located at $R_i$. Accurate calculations of the entire energy surface $E(R_i)$ are possible with traditional (wave function based) methods only for systems with few atoms, and $E(R_i)$ generally has vast numbers of maxima and minima at unknown locations.
The lowest energy, however, corresponds to the ground state structure, and paths between minima are essential to studies of chemical reactions, including their activation energies. The observation of Francis Crick in his autobiography (Crick, 1990, p. 150): “If you want to study function, study structure” may be self-evident to anyone interested in biology or molecules in general, but it is true in other areas. The DF approach allows us to calculate $E(R_1)$ and hence the structure and many related properties, without using experimental input.

Olle Gunnarsson and I reviewed the DF formalism, its history, and its prospects in 1989 (Jones and Gunnarsson, 1989). A careful reading of the original literature some years ago suggested an alternative approach, and I trace here the DF history from the first years after the development of quantum mechanics. It is a fascinating story with many players, and I quote in several places from the original texts.

Density functional calculations are now well established in condensed matter physics and chemistry, but they did not (and do not) find universal acceptance. The choice of 1990 as fixed point coincides with the publication of the review of Jones and Gunnarsson (1989), the book of Parr and Yang (1989), and an article (Jones, 1991) advocating DF calculations for molecules, particularly when combined with molecular dynamics (MD) (Car and Parrinello, 1985). Figure 1 (Mavropoulos, 2015) shows that it also marks the dramatic increase in the number of publications on the topics “density functional” and density functional theory (“DFT”) in recent years. The relatively small number of publications before 1990 by no means implies that important work was not being carried out. Many applications that today would be denoted “density functional” then used other designations.

In this article, I revisit the period before 1990 [some aspects of which are covered in detail by Parr and Yang (1989)] and focus on developments since then. Perspectives on density functional theory have been given by Burke (2012) and Becke (2014), and much more detailed information is available in monographs and review articles cited there. A recent issue of the Journal of Chemical Physics celebrated 50 years of modern density functional theory, and the articles range across many topics of current interest (Yang, 2014). A review of solid state applications of DF theory is provided by Hasnip et al. (2014), and Zangwill (2014) discussed the life and work of Walter Kohn, particularly in this context. I use numerous citations of participants in (and observers of) this story to illustrate how perspectives on the approach have changed. The two applications I discuss are DF simulations of systems that were unimaginable only a few years ago.

I have several goals in writing this review. The large growth in the DF literature is possible only if there are many recent arrivals in the field. This review is written for them and for those in other areas of science who are curious about the DF world. I hope that all share my fascination with the formalism and its history, sense the excitement that being part of a developing field can bring, and appreciate that scientific research is carried out by people, not machines. Personal contacts over the years have shown, however, that it is not only newcomers who are unfamiliar with the past or the reaction of different scientific communities as the theory developed. Density-related methods are also important in other areas, including classical systems and nuclei, and I encourage interested readers to look beyond the horizons of their particular interest. Finally, I note that both Burke (2012) and Becke (2014) are uneasy about some recent developments, and I shall raise my own questions about the future.

II. THE DENSITY AS A BASIC VARIABLE

The recent books by Segrè (2007) and Farmelo (2009) give fascinating accounts of the development of quantum mechanics in the years following 1926. Methods for finding approximate solutions of the Schrödinger equation followed soon after the equations were published and have had a profound effect on chemistry and condensed matter physics ever since.

A method for calculating the wave function of an atom was developed by Hartree (1928a, 1928b), who introduced the idea of a “self-consistent field” with particular reference to valence electrons and groups of core electrons. In this approach, the wave function of an electron $\psi_i$ is determined from the field of the nucleus and the other electrons in a self-consistent fashion. One starts with an approximate field [such as one derived using the Thomas-Fermi (TF) approximation discussed later] and iterates until input and output fields for all electrons are the same.

The wave function of the $N$-electron system can be approximated by the product of $N$ single-particle functions,

$$\Psi(r_1, r_2, \ldots) = \psi_1(r_1) \cdots \psi_N(r_N),$$

\footnote{A similar plot is given by Burke (2012) for two popular approximations used in DF calculations.}
where each \( \psi_i(\mathbf{r}) \) satisfies a one-electron Schrödinger equation with a potential term arising from the average field of the other electrons. Equation (1) is often termed the “Hartree approximation,” but it is not mentioned in either paper of Hartree cited previously. Fock (1930) and Slater (1930) recognized that the product wave function (1) in conjunction with the variational principle led to a generalization of the method that would apply to systems more complex than atoms. Fock (1930) and Slater (1930) also showed that replacing Eq. (1) by a determinant of such functions led to equations that were only a little more complicated, while satisfying the Pauli exclusion principle. These determinantal functions, which had been used in discussions of atoms (Slater, 1929) and ferromagnetism (Bloch, 1929), are known today as “Slater determinants,” and the resulting “Hartree-Fock equations” have formed the basis of most discussions of atomic and molecular structure since.

Just three years after the derivation of the Schrödinger equation, Dirac (1929a) wrote

“The general theory of quantum mechanics is now almost complete, … The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble. It therefore becomes desirable that approximate practical methods of applying quantum mechanics should be developed, which can lead to an explanation of the main features of complex atomic systems without too much computation.”

Dirac emphasizes the difficulty of solving the equations of quantum mechanics and the desirability of developing “approximate practical methods of applying quantum mechanics” to explain complex systems. I cannot think of a better short description of the motivation of density functional theory.

The first “density functional theory” for electronic systems was given by Thomas (1927) and Fermi (1927, 1928), who described a model for calculating atomic properties based purely on the electron density \( n(\mathbf{r}) \).\(^4\) They assumed that the electrons form a gas satisfying Fermi statistics, with the electron-electron interaction energy determined from the classical Coulomb potential. For the kinetic energy they adopted a local density (LD) approximation, where the contribution from the point \( \mathbf{r} \) [where the density is \( n(\mathbf{r}) \)] is determined from the kinetic energy of a homogeneous electron gas with this density.\(^5\) The TF Euler equation for the density is

\[
\frac{5}{3} C_N n(\mathbf{r})^{2/3} + e^2 \int d\mathbf{r'} \frac{n(\mathbf{r'})}{|\mathbf{r} - \mathbf{r'}|} + V_{\text{ext}}(\mathbf{r}) + \lambda = 0, \tag{2}
\]

where \( C_N = 3\hbar^2/(3\pi^2)^{2/3}/(10n) \), \( V_{\text{ext}} \) is the external potential, and \( \lambda \) is the Lagrange multiplier related to the constraint of constant particle number.

The Thomas-Fermi method and its extensions approximately describe the charge density, the electrostatic potential, and the variation of the total energy with atomic number \( Z \), and its mathematical properties have attracted considerable attention (Schwinger, 1980; Lieb, 1981; Spruch, 1991). Although the charge density is infinite at the nucleus, the leading term in the expansion of the energy as a function of atomic number \( Z \) is exact for atoms and periodic solids in the limit of large \( Z \) (Lieb and Simon, 1973). This limit describes the “inner core” and “mantle” regions of an atom, both of which shrink (as \( Z^{-1/3} \)) with increasing \( Z \) (Lieb and Simon, 1973). These properties of the TF approach at large \( Z \) mean that it has been applied to stellar and other matter at extremely high pressures and densities.\(^6\)

Nevertheless, TF theory has severe deficiencies because of its poor description of the outer regions of an atom. The charge density decays as \( r^{-6} \) far from the nucleus, not exponentially as it should. Teller (1962) showed also that TF theory does not bind neutral atoms or (with some restrictions) ions to form molecules or solids, which rules out its use in chemistry or materials science at normal temperatures and pressures. The local density approximation for the kinetic energy means that there is also no shell structure in the TF atom, so that the periodic variation of many properties with changing atomic number \( Z \) cannot be reproduced. Furthermore, the lowest energy of a system has zero spin, so that ferromagnetism cannot occur (Jones and Gunnarsson, 1989). Nordholm (1987) has traced the inability of the TF approach to describe bonding to the assumption of ergodicity, i.e., that the electron density is able to fill all available phase space up to the Fermi energy, whether or not these parts of phase space are dynamically connected. This assumption constrains dynamical processes in atoms and molecules.

Dirac (1930a) noted the necessity of incorporating “exchange” phenomena into the “Thomas atom,” and he did this by recasting Hartree-Fock theory in terms of a “density function,” without reference to a single-determinant many-electron wave function. In the context of the Thomas-Fermi theory, this function leads to a correction to the energy derived from the exchange energy for a homogeneous electron gas of density \( n \).\(^7\) The corresponding potential is

\[
V_{\chi}^{\text{Dirac}} = -\left(\frac{1}{16} \right) [3\pi^2 n(\mathbf{r})]^{1/3}. \tag{3}
\]

The modified TF equation is often referred to as the “Thomas-Fermi-Dirac” equation.

\(^4\)The roles of Hartree and Thomas as precursors of modern density functional theory are discussed by Zangwill (2013).

\(^5\)Fermi later extended the model to positive ions and spectroscopic energy levels. It was pointed out (Guerra and Robotti, 2008) that the extension to positive ions and to Rydberg corrections was described by Majorana already in 1928 [an English translation is provided in Guerra and Robotti (2008)].

\(^6\)See, for example, Spruch (1991). Smith et al. (2014) used TF theory to analyze data on ramp compressed diamond to 5 TPa.

\(^7\)The exchange energy in a homogeneous electron gas had already been derived by Bloch (1929).
One point made by Dirac (1930a) has been emphasized by many advocates of the DF method over the years, even if we were unaware of his words of over 80 years ago:

“Each three-dimensional wave function will give rise to a certain electric density. This electric density is really a matrix, like all dynamical variables in the quantum theory. By adding the electric densities from all the wave functions we can obtain the total electric density for the atom. If we adopt the equations of the self-consistent field as amended for exchange, then this total electric density (the matrix) has one important property, namely, if the value of the total electric density at any time is given, then its value at any later time is determined by the equations of motion. This means that the whole state of the atom is completely determined by this electric density; it is not necessary to specify the individual three-dimensional wave functions that make up the total electric density. Thus one can deal with any number of electrons by working with just one matrix density function.”

The italics are in the original. The term “density function” is ambiguous, but Dirac had defined here what we know today as the one-particle reduced density matrix \( \rho(\mathbf{r}, \mathbf{r'}) \), which is related to the density by \( n(\mathbf{r}) = \rho(\mathbf{r}, \mathbf{r}) \) and is discussed further in Sec. VI.A.6. The observation that the density follows the equations of motion is much in the spirit of the theorem of Ehrenfest (1927), who proved what has been termed the “time-dependent Hellmann-Feynman theorem” (Hayes and Parr, 1965), namely, that the acceleration of a quantum wave packet that does not spread satisfies Newton’s equations of motion.

The central role played by the electron density means that it is essential to have a clear picture of its nature in real systems. In Fig. 2, we show the spherically averaged density in the ground state of the carbon atom. The density falls monotonically from the nucleus and does not show the radial oscillations that occur if one were to plot \( r^2 n(r) \). The charge density in small molecules is also relatively featureless, with maxima at the nuclei, saddle points along the bonds, and a generally monotonic decay from both. Electron densities in molecules and solids also show relatively small departures from the overlapped densities of the constituent atoms. Energy differences, including binding, ionization, and cohesive energies, are the focus of much DF work and result from subtle changes in relatively featureless density distributions. It really is astonishing that these distributions suffice to determine ground state properties and, in principle, excited state properties as well.

\[ 8 \quad \text{See also Dirac (1929b, 1930b). Dirac was not the only one to use different expressions or even different definitions for the density matrix. For a discussion, see ter Haar (1960).} \]

III. AN “APPROXIMATE PRACTICAL METHOD”

The basis of a quantum theory of atoms, molecules, and solids was in place at the beginning of the 1930s. Linear combinations of atomic orbitals formed molecular orbitals, from which determinantal functions could be constructed, and linear combinations of determinants (“configuration interaction,” CI) would provide approximations to the complete wave function. Dirac had noted already, however, that this procedure could not be implemented in practice, so that approximations are essential. Furthermore, numerical techniques for solving the Schrödinger equation in extended systems needed to be developed.

A. Exchange hole and local density approximations

Wigner and Seitz (1933, 1934) developed a method for treating the wave function in crystals, and the “Wigner-Seitz cell” construction is known to all condensed matter physicists. The first application to metallic sodium replaced the nucleus and core electrons by an effective (pseudo)potential, and calculations of the lattice constant, cohesive energy, and compressibility gave satisfactory results. Of particular interest to us is the calculation of the probability of finding electrons with parallel spin components a distance \( r \) apart (Fig. 3). This function obtains its half value for \( r = 1.79d \) or \( 0.460d \) for a body-centered cubic lattice with cube edge \( d \), which is close to the radius of the sphere whose volume is the mean volume per atom of a solid (the “Wigner-Seitz sphere”), \( (3/8\pi)^{1/3} d = 0.492d \). The exclusion principle means then that two electrons with parallel spins will usually be associated with different ions (Wigner and Seitz, 1933). The corresponding curves for spin-up and spin-down electrons, as well as for both spins combined, were discussed by Slater (1934). Similar views of electrons in metals have been given by Brillouin (1934) and Wigner (1934).
FIG. 3. Probability that electrons with parallel spins are \(r/d'\) apart in Na metal \((d' = V_0/3\pi^2, V_0\) is the atomic volume). Adapted from Wigner and Seitz, 1933.

The picture that results is simple and appealing: The exclusion principle means that an electron with a given spin produces a surrounding region where there is a deficiency of charge of the same spin. This region contains unit charge and is referred to as the Fermi (Wigner and Seitz, 1933) or exchange hole (Slater, 1951). This hole is of crucial importance to DF theory and plays a central role in the discussion to follow. In the Hartree-Fock scheme, the exchange hole is different for each electronic function, but Slater (1951) developed a simplified “exchange potential” that depended only on the density\(^9\)

\[
V_{\text{Slater}} = -\left(\frac{3}{2\pi}\right) [3\pi^2 n(\mathbf{r})]^{1/3}. \tag{4}
\]

Sharp and Horton (1953) also noted the advantages of an effective potential that is the same for all electrons. They constructed an “optimized effective potential” (OEP) by varying the potential in a one-body Hamiltonian to produce orbitals for a Slater determinant that minimizes the expectation value of the many-electron Hamiltonian. Talman and Shadwick (1976) showed that this effective potential has the correct asymptote far from an atom. The OEP is often discussed in the context of DF calculations, but it was originally an independent development.

The Slater approximation [Eq. (4)] was proposed at the time that electronic computers were becoming available for electronic structure calculations and proved to be very useful in practice. Methods for solving the Schrödinger equation had been developed around this time, including the augmented plane-wave (APW) (Slater, 1937) and Korringa-Kohn-Rostoker approaches (Korringa, 1947; Kohn and Rostoker, 1954).

The exchange potential of Slater is \(3/2\) times that derived by Bloch (1929) and Dirac (1930a) for a homogeneous electron gas [Eq. (3)], but Slater (1968) pointed out that an effective potential proportional to the cube root of the density could be obtained by arguments based on the exchange hole that are independent of the homogeneous electron gas arguments used in the original derivation (Slater, 1951). The exchange hole discussed for a spin-up electron contains a single electron. If we assume that the hole can be approximated by a sphere of radius \(R_\uparrow\), then

\[
\left(\frac{4\pi}{3}\right) R_\uparrow^3 n_\uparrow = 1, \quad R_\uparrow = \left(\frac{3}{4\pi n_\uparrow}\right)^{1/3}, \tag{5}
\]

where \(n_\uparrow\) is the density of spin-up electrons. Since the electrostatic potential at the center of such a spherical charge is proportional to \(1/R_\uparrow\), the exchange potential will be proportional to \(n_\uparrow^{1/3}\). This argument was used by Slater (1972b) to counter a (still widespread) misconception that local density approximations are valid only if the electron density is nearly homogeneous.

Gáspár (1954) questioned the prefactor of the effective exchange potential [Eq. (4)]. If one varies the spin orbitals to minimize the total energy in the Thomas-Fermi-Dirac form, one obtains a coefficient just \(2/3\) as large. Gáspár applied this approximation to the Cu\(^+\) ion and found good agreement with Hartree-Fock eigenfunctions and eigenvalues. Slater noted that Gáspár’s method was “more reasonable than mine” (Slater, 1974), and he adopted the procedure in his later work (Slater, 1972a, p. 23):

“There are important advantages in the Gáspár-Kohn-Sham procedure of using a statistical expression for total energy and varying the orbital to minimize the energy, and we shall henceforth adopt this approach.”

However, the original APW code and its early applications used the Slater potential, which was modified by an adjustable parameter \(\alpha\) \((\alpha = 2/3\) gives the exchange energy of a homogeneous electron gas), and the approximation came to be known as the “\(X\alpha\) method” (see Sec. IV.B).\(^10\)

B. Electron density and the chemical bond

Slater and Gáspár were not alone in their focus on the density. The “Hellmann-Feynman theorem” (Hellmann, 1933,

\(^9\)Slater wrote in this paper: “The discussion of Wigner and Seitz was one of the first to show a proper understanding of the main points taken up in this paper, which must be understood to represent a generalization and extension of previously suggested ideas, rather than an entirely new approach. See also Brillouin (1934) for a discussion similar to the present one.” Slater was more generous to his predecessors than some were to him later. See, for example, Anderson (2011), pp. 125–127) and Kohn and Sham (1965).

\(^{10}\)In a conference discussion, Kohn (1971) noted: “The correlation correction has a completely different density dependence (from exchange), and the rational way to handle it is to do it right. In the very high density limit it becomes negligible compared to the exchange. It does not go with the same power of the density, and so certainly the way not to handle it is to just put a constant in front and say it is somewhere between 2/3 and 1.”
1937; Feynman, 1939) (see the Appendix)\(^{11}\) considered forces in molecules, and Berlin (1951) developed this picture to separate space near a diatomic molecule into “binding” and “antibinding” regions according to the sign of the electrostatic interaction between the nuclei and the local electron charge distribution. The Hellmann-Feynman equations are exact only for the exact wave function of the system. In his “electrostatic method,” however, Hurley (1954a, 1954b, 1954c) showed that force calculations could be performed with approximate wave functions if all variable parameters in the wave function for a molecule were chosen variationally. The procedure is simplified for “floating functions,” where the parameters do not depend on the nuclear configurations.\(^{12}\) This approach has proved to be fruitful in many contexts, including dispersion interactions (Strømsheim et al., 2011). It was around this time that Bader (1990) began his studies of the topology of the density distribution that he pursued for many years.

Electrostatic arguments are often presented as the main cause of chemical bonding: a molecule has a lower potential energy than its component atoms because the increased electron density in the interatomic region increases the effective internuclear attraction. The virial theorem also tells us that the lowering of the potential energy on moving two atoms from infinite separation to the equilibrium distance is twice as much as the increase in kinetic energy. Can the chemical bond really be viewed in such purely classical terms?

An early argument against this picture was given in the original paper of Hellmann (1933), who noted that the delocalization of the valence electron wave functions would lower the kinetic energy. This argument was extended by Ruedenberg (1962), who emphasized that electron sharing (delocalization) lowered the potential energy as well.\(^{13}\) Calculations of \(\text{H}_2^+\) and \(\text{H}_2\) (Ruedenberg, 1962; Kutzelnigg and Schwarz, 1982; Bacsay and Nordholm, 2013) enabled different bonding components to be isolated and showed that bond formation is also accompanied by orbital contraction. Ruedenberg viewed the formation of a covalent bond as a two-step process: (1) a contraction of the atomic density prior to bonding, and (2) a lowering of the kinetic energy as the atoms are brought to their equilibrium separation. Kutzelnigg (1973) reversed the order of these steps in a discussion of \(\text{H}_2^+\), where the latter is significantly more important than the ultimate contraction of the atomic densities, which simply adjusts the balance between kinetic and potential energies to satisfy the virial theorem. This picture of the physical origin of the chemical bond is accepted by many theoretical chemists, but it can be found in few chemistry textbooks.

A variant of the Hellmann-Feynman theorem with a variable nuclear charge \(Z\) was used by Wilson (1962) to show that the ground state energy of a system of electrons in the field of a set of fixed nuclei can be calculated if we know the electron density sufficiently accurately as a function of the spatial coordinates \(x, y, z\), and the parameter \(Z\) from 0 (noninteracting system) to 1 (physical system). He concluded with the following challenge:

> “The important question remains: Does there exist some procedure for calculating (the density) \(n\) which avoids altogether the use of \(3N\)-dimensional space? Such a procedure might open the way to an enormous simplification of molecular calculations. For example, it would be tremendously simpler to expand a four-dimensional function than a \(3N\)-dimensional wave function.”

He did not have to wait long for an answer.

### IV. Modern Density Functional Formalism

The variational principle on the energy was the basis of the formulation of the exact density functional formalism given by Hohenberg and Kohn (1964) (HK). First they showed that there is a one-to-one relationship between the external potential \(V_{\text{ext}}(\mathbf{r})\) and the (nondegenerate) GS wave function \(\Psi\), and that there is a one-to-one relationship between \(\Psi\) and the ground state density \(n(\mathbf{r})\) of an \(N\)-electron system.\(^{14}\)

\[
n(\mathbf{r}) = N \int d\mathbf{r}_2 \cdots d\mathbf{r}_N \Psi^*(\mathbf{r}, \mathbf{r}_2, \ldots, \mathbf{r}_N)\Psi(\mathbf{r}, \mathbf{r}_2, \ldots, \mathbf{r}_N),
\]

where the spin coordinates are not shown explicitly. Knowledge of the density then determines the external potential to within a constant, so that all terms in the Hamiltonian are known. Since the Hamiltonian operator completely determines all states of the system, \(n(\mathbf{r})\) determines excited states as well as the ground state.\(^{15}\)

These ideas can be applied to the total energy using the variational principle. For this purpose, HK defined the functional \(F[n(\mathbf{r})]\), which is “universal” in the sense that it is valid for any external potential \(V_{\text{ext}}\),

\[
F[n] = \langle \Psi_n | T + V_{\text{ee}} | \Psi_n \rangle,
\]

and showed that the energy functional \(E[n, V_{\text{ext}}]\) satisfies a variational principle:

\[
E_{\text{GS}} = \min_{n(\mathbf{r})} E[n, V_{\text{ext}}],
\]

\(^{11}\)An interesting account of the history of the Hellmann-Feynman theorem and its mathematical complexities is given in Pupyshev (2000).

\(^{12}\)See also Frost (1967).

\(^{13}\)For a recent survey see Bacsay and Nordholm (2013).

\(^{14}\)The relationship between electron density and \(\Psi^*\Psi\) was first discussed by Schrödinger (1926b). Shortly afterward, Born (1926a) introduced the concept of “probability” in the context of electron-atom scattering.

\(^{15}\)At a conference in 1965, E. B. Wilson noted (Handy, 2009): If we know the density distribution, we can find its cusps, from which we can determine [from the electron-nucleus cusp condition (Kato, 1957)] both the location of the nuclei and their atomic numbers. Integration over the electron density gives the total charge, so that we have all the terms in the Hamiltonian. This nonrelativistic argument applies for a system with Coulomb interactions and assumes point nuclei and the Born-Oppenheimer approximation.
where

\[ E[n, V_{\text{ext}}] = \int d\mathbf{r} V_{\text{ext}}(\mathbf{r}) n(\mathbf{r}) + F[n]. \]  

(9)

The minimization is performed in HK over all nondegenerate densities that can be derived from the ground state of some external potential (V-representable). The generalization of Levy (1979) to minimizing over all densities, including degeneracies, is discussed in Sec. V.A.2.

A. Single-particle description of a many-electron system

The task of finding good approximations to the energy functional \( E[n, V_{\text{ext}}] \) is greatly simplified if we use the decomposition introduced by Kohn and Sham (1965)\(^{16}\):

\[ F[n] = T_s[n] + \frac{1}{2} \int d\mathbf{r} n(\mathbf{r}) \Phi(\mathbf{r}) + E_{xc}[n]. \]  

(10)

\( T_s \) is the kinetic energy that a system with density \( n \) would have in the absence of electron-electron interactions, \( \Phi \) is the classical Coulomb potential for electrons, and \( E_{xc} \) defines the exchange-correlation (xc) energy.\(^\text{17}\) \( T_s \) is not the true kinetic energy \( T \), but it is of comparable magnitude and is treated here without approximation. This removes many of the deficiencies of the Thomas-Fermi approach, such as the lack of a shell structure of atoms or the absence of chemical bonding in molecules and solids. All terms in Eq. (10) other than the exchange-correlation energy \( E_{xc} \) can be evaluated exactly, so that approximations for this term are crucial in density functional applications.

The variational principle applied to Eq. (10) yields

\[ \frac{\delta E[n, V_{\text{ext}}]}{\delta n(\mathbf{r})} = \frac{\delta T_s}{\delta n(\mathbf{r})} + V_{\text{ext}}(\mathbf{r}) + \Phi(\mathbf{r}) + \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})} = \mu, \]  

(11)

where \( \mu \) is the Lagrange multiplier associated with the requirement of constant particle number. If we compare this with the corresponding equation for a system with the same density in an external potential \( V(\mathbf{r}) \) but without electron-electron interactions,

\[ \frac{\delta E[n]}{\delta n(\mathbf{r})} = \frac{\delta T_s}{\delta n(\mathbf{r})} + V(\mathbf{r}) = \mu, \]  

(12)

we see that the mathematical problems are identical, provided that

\[ V(\mathbf{r}) = V_{\text{ext}}(\mathbf{r}) + \Phi(\mathbf{r}) + \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})}. \]  

(13)

The solution of Eq. (12) can be found by solving the Schrödinger equation for noninteracting particles.

\[ [-\frac{1}{2} \nabla^2 + V(\mathbf{r})] \phi_i(\mathbf{r}) = c_i \phi_i(\mathbf{r}), \]  

(14)

yielding

\[ n(\mathbf{r}) = \sum_{i=1}^{N} f_i |\phi_i(\mathbf{r})|^2. \]  

(15)

The functions \( \phi_i \) are the Kohn-Sham (KS) orbitals, and the occupation numbers \( f_i \) are noninteger at zero temperature when the orbitals are degenerate at the Fermi level and Fermi-Dirac occupancies at nonzero temperatures. The condition (13) can be satisfied in a self-consistent procedure.

The solution of this system of equations leads to the energy and density of the lowest state and all quantities derivable from them. The formalism can be generalized to the lowest state with a given symmetry (Gunnarsson and Lundqvist, 1976) or other constraints (Dederichs et al., 1984). Instead of seeking these quantities by determining the wave function of the system of interacting electrons, the DF method reduces the problem to the solution of a single-particle equation of Hartree form. In contrast to the nonlocal Hartree-Fock potential,

\[ V_{\text{HF}}(\mathbf{r}) = \int d\mathbf{r}' V_{\text{HF}}(\mathbf{r}, \mathbf{r}') \psi(\mathbf{r}'), \]  

(16)

the effective potential \( V(\mathbf{r}) \) is a local (multiplicative) operator.

The numerical advantages of solving the Kohn-Sham equations (Kohn and Sham, 1965) are obvious. Efficient methods exist for solving (self-consistently) single-particle Schrödinger-like equations with a local effective potential, and there is no restriction to small systems. With a local approximation to \( E_{xc} \), the equations can be solved as readily as the Hartree equations. Unlike the Thomas-Fermi method, where the large kinetic energy term is computed directly from the density, this term is computed exactly from orbitals \( \phi_i \) in Eq. (14). The core-valence and valence-valence electrostatic interactions can be directly evaluated, but \( E_{xc} \) is the difference between the exact energy and terms we can accurately evaluate, and approximations are unavoidable.

B. Exchange-correlation energy \( E_{xc} \)

Kohn and Sham (1965) proposed using the LD approximation

\[ E_{xc}^{LD} = \int d\mathbf{r} n(\mathbf{r}) \epsilon_{xc}[n(\mathbf{r})], \]  

(17)

where \( \epsilon_{xc}[n] \) is the exchange and correlation energy per particle of a homogeneous electron gas with density \( n \). This is a good approximation if the density is almost constant (Hohenberg and Kohn, 1964), as well as at high densities, where the kinetic energy dominates the exchange and correlation terms (Kohn and Sham, 1965). They noted that this approximation “has no validity” at the “surface” of atoms and in the overlap regions of molecules and concluded (Kohn and Sham, 1965).
“We do not expect an accurate description of chemical bonding.”

The applicability of this comment depends on the definition of “accurate,” but this view was shared by many and certainly delayed chemical applications of DF theory by some years. It was a legacy of the origins of the LD approximation in the homogeneous electron gas and ignored the exchange-hole argument of Slater noted in Sec. III.A.

Generalizations to spin-polarized systems were given by von Barth and Hedin (1972) and Rajagopal and Callaway (1973). The local spin density (LSD) approximation can be written

\[ E_{\text{LSD}} = \int d\mathbf{r} n(\mathbf{r}) \varepsilon_{\text{xc}}(n_\uparrow(\mathbf{r}), n_\downarrow(\mathbf{r})), \quad (18) \]

where \( \varepsilon_{\text{xc}}[n_\uparrow, n_\downarrow] \) is the exchange and correlation energy per particle of a homogeneous, spin-polarized electron gas with spin-up and spin-down densities \( n_\uparrow \) and \( n_\downarrow \), respectively (von Barth and Hedin, 1972).\(^\text{18}\) The \( X\alpha \) approximation

\[ E_{\alpha\alpha} = -\frac{3}{2} a C \int d\mathbf{r} \left\{ [n_\uparrow(\mathbf{r})]^{4/3} + [n_\downarrow(\mathbf{r})]^{4/3} \right\}, \quad (19) \]

where \( C = 3/(3/4\pi)^{1/3} \) was used in numerous calculations in the late 1960s and 1970s. The \( \alpha \) dependence of energy differences for a given atom or molecule is weak for values near \( 2/3 \), the value of Bloch (1929), Dirac (1930a), Gáspár (1954), and Kohn and Sham (1965). We noted that the electron density in molecules and solids is generally far from that of a homogeneous electron gas and that the validity of calculations based on properties of a gas of constant density has often been questioned. We now discuss some general properties of \( E_{\alpha\alpha} \) using arguments closely related to the “exchange-hole” picture of Wigner and Seitz (1933) and Slater (1934, 1951, 1968).

C. Exchange-correlation hole and \( E_{\alpha\alpha} \)

The crucial simplification in the density functional scheme is the relationship between the interacting system, whose energy and density we seek, and the artificial (sometimes called fictitious), noninteracting system for which we solve Eqs. (14) and (15). This can be studied by considering the interaction \( \lambda/|\mathbf{r} - \mathbf{r}'| \) and varying \( \lambda \) from 0 (noninteracting system) to 1 (physical system).

Two ways of doing this have proved useful in the DF context. The first (Harris and Jones, 1974) considered the Hamiltonian

\[ H_\lambda = T + V + \lambda(V_{\text{ee}} + V_{\text{ext}} - V), \quad (20) \]

where \( V \) is the (density-dependent) Kohn-Sham effective potential [Eq. (13)] and \( V_{\text{ee}} \) is the electron-electron interaction. The density for \( \lambda = 0 \) is constrained to be that of the physical system, and the qualitative difference between the eigenfunctions for \( \lambda = 0 \) and \( \lambda = 1 \) is reduced. This construction leads to an expression for the total energy \( E \) and its components, including the exchange and correlation energies, in terms of integrals over \( \lambda.\(^\text{19}\) One term can be related to the energy due to density fluctuations in the system and the density response function of a system of electrons described by \( H_\lambda \), and all terms were evaluated for a model of a bounded electron gas (Harris and Jones, 1974).\(^\text{20}\)

This approach, in which the density is constrained to be the physical density of the system for \( \lambda = 0 \) and 1, but not for intermediate values, can be implemented in standard programs for calculating wave functions and provides detailed information about the density functional in small systems (Savin, Colonna, and Teuler, 1998). It is possible, however, to adopt an external potential \( V_\lambda \) such that the ground state of the Hamiltonian \( H_\lambda \) has density \( n(\mathbf{r}) \) for all \( \lambda \) (Langreth and Perdew, 1975; Gunnarsson and Lundqvist, 1976).\(^\text{21}\) The exchange-correlation energy of the interacting system can then be expressed as an integral over the coupling constant \( \lambda \) (Langreth and Perdew, 1975; Gunnarsson and Lundqvist, 1976):

\[ E_{\alpha\alpha} = \frac{1}{2} \int d\mathbf{r} n(\mathbf{r}) \int d\mathbf{r}' \frac{1}{|\mathbf{r} - \mathbf{r}'|} n_{\alpha\alpha}(\mathbf{r}, \mathbf{r}' - \mathbf{r}), \quad (21) \]

with

\[ n_{\alpha\alpha}(\mathbf{r}, \mathbf{r}' - \mathbf{r}) = n(\mathbf{r}') \int_0^1 d\lambda [g(\mathbf{r}, \mathbf{r}', \lambda) - 1]. \quad (22) \]

The form of this expression, with the physical \( n(\mathbf{r}) \) occurring as a term in the integration over space, is convenient in the DF context [cf. Eqs. (17) and (18)] and is more widely used.

The function \( g(\mathbf{r}, \mathbf{r}', \lambda) \) is the pair-correlation function of the system with density \( n(\mathbf{r}) \) and Coulomb interaction \( \lambda V_{\text{ee}} \). The exchange-correlation hole \( n_{\alpha\alpha} \) describes the fact that an electron at point \( \mathbf{r} \) reduces the probability of finding one at \( \mathbf{r}' \), and \( E_{\alpha\alpha} \) is simply the energy resulting from the interaction between an electron and its exchange-correlation hole. This is a straightforward generalization of the work of Wigner and Seitz (1933, 1934) and Slater (1934, 1951) discussed in Sec. III.A.

---

\(^\text{18}\) This appears to be the first use of “adiabatic connection” in the DF context. The derivative of the Hamiltonian with respect to a parameter and the integral form that follows are often attributed to Pauli (1933) [see Musher (1966) and references therein]. Pauli (1933, p. 161) gives the credit to his student Güttinger (1931). See the Appendix.

\(^\text{20}\) This calculation was motivated in part by our expectation that the LD approximation would not describe well the surface energy of a bounded electron gas. Our estimates of the exchange and correlation contributions were individually different from the LD results, but the sum of the two components was similar for the values of the bulk densities we calculated. This unexpected success of the LD approximation (for exchange and correlation) suggested that we try to understand what was behind it. Quantum Monte Carlo calculations confirmed the LD results many years later (Wood et al., 2007).

\(^\text{21}\) It has been shown (Colonna and Savin, 1999) that results for the He and Be series atoms using the two schemes are similar.
This means that we can consider atom is shown for a representative value of for both the local and does not depend on the nonspherical parts of moment of a function whose second moment we know exactly. Third, the definition of the pair-correlation function leads to a sum rule requiring that the xc hole corresponds to the removal of one electron, i.e., for all \( r \),

\[
\int dr' n_{xc}(r, r' - r) = -1. \tag{24}
\]

This means that we can consider \(-n_{xc}(r, r' - r)\) as a normalized weight factor, and define the radius of the xc hole locally for a particular value of \( r \):

\[
\left\langle \frac{1}{R} \right\rangle_r = -\int dR \frac{n_{xc}(r, R)}{|R|}. \tag{25}
\]

This leads to

\[
E_{xc} = -\frac{1}{2} \int dr n(r) \left\langle \frac{1}{R} \right\rangle_r. \tag{26}
\]

Provided Eq. (24) is satisfied, \( E_{xc} \) is determined by the first moment of a function whose second moment we know exactly and does not depend on the nonspherical parts of \( n_{xc} \) (Gunnarsson and Lundqvist, 1976). Approximations to \( E_{xc} \) can lead to good total energies (and structures) for density distributions far from the regions of obvious validity. This is illustrated in Fig. 4, where the exchange hole in a nitrogen atom is shown for a representative value of \( r \) for both the local density and exact (Hartree-Fock) cases. The holes are qualitatively different: The LD hole is spherically symmetric and centered on the electron, while the exact hole has a large weight at the nucleus and is asymmetric. Nevertheless, the spherical averages are similar, and this is true for other values of \( r \). The exchange energies found from Eq. (26) differ by just over 10% (LSD: \(-11.79\) Ry, exact: \(-13.19\) Ry) (Gunnarsson, Jonson, and Lundqvist, 1979).

V. ADVANCES TO 1990

A. Theoretical progress

The variation in the number of citations to density functional theory and related topics (Fig. 1) should be interpreted with caution. The dramatic increase after 1990 hides the fact that many articles that we associate today with density functional or DFT (Hartree-Fock-Slater, Xα, electronic band structure) made no reference to it. It seems that many realized around this time that they had been doing density functional calculations all along. However, theoretical and computational work was already in place by 1990 that has proved to be crucial to the ultimate acceptance of the method, and I now outline some of it. More details can be found in Jones and Gunnarsson (1989), Parr and Yang (1989), and Dreizler and Gross (1990).

1. Temperature, spin, and time dependence

The generalizations to finite temperatures and to spin systems were carried out soon after the original work of Hohenberg and Kohn (1964). The former was provided by Mermin (1965), who showed that, in a grand canonical ensemble at given temperature \( T \) and chemical potential \( \mu \), the equilibrium density is determined by the external potential \( V_{ext} \), and the equilibrium density minimizes the grand potential. Single-particle equations can be derived for an artificial system with kinetic energy \( T_\alpha \) and entropy \( S_\alpha \), with \( E_{xc} \) replaced by the exchange-correlation contribution to the free energy.

The extension to spin systems (von Barth and Hedin, 1972; Rajagopal and Callaway, 1973) or an external magnetic field requires the introduction of the spin indices \( \alpha \) of the one-electron operators \( \psi_\alpha (r) \) and replacing \( V_{ext} \) by \( V^{\alpha \beta}_{ext}(r) \) and
the charge density $n(r)$ by the density matrix $n_{\alpha\beta}(r) = \langle \Psi | \phi_{\alpha}(r) \phi_{\beta}(r) | \Psi \rangle$. All ground state properties are functionals of $n_{\alpha\beta}$, and $E$ is stationary with respect to variations in $n_{\alpha\beta}$. The expression for the energy $E_{xc}$ is analogous to Eqs. (21) and (22). A current and spin density functional theory of electronic systems in strong magnetic fields was formulated by Vignale and Rasolt (1988). Time-dependent density functional theory, which has proved to be valuable in discussing excited states, was described by Mahan (1979), Stott and Zaremba (1980), and Zangwill and Soven (1980), who introduced the adiabatic local density approximation. Runge and Gross (1984) proved the time-dependent analog of the Hohenberg-Kohn theorem.

2. Developments of formalism

In the spirit of the definition by Percus (1978) of a universal kinetic energy functional for noninteracting systems, Levy (1979) provided a proof that was simpler and more general than that of Hohenberg and Kohn (1964). The sum of the kinetic and electron-electron repulsion energies [Eq. (9)] is written

$$F[n] = \min_{\Psi = n(r)} \langle \Psi | T + V_{ee} | \Psi \rangle,$$

where the minimization is performed over all antisymmetric wave functions $\Psi$ that lead to the density $n(r)$. Harriman (1981) showed that these functions exist, since all mathematically well-behaved densities can be obtained from antisymmetric wave functions. The removal of the Hohenberg-Kohn restrictions to nondegenerate states and densities that could be determined from the ground state of some effective potential was an important development and enabled DF practitioners to avoid some of the concerns then current. These “constrained search” arguments were extended by Lieb (1983), who also introduced the Legendre transform of the energy as the universal functional. The use of the variational principle in these works goes beyond the insights of Dirac (1930a) and Wilson (1962).

Important developments during this period included work on constraints that must be satisfied by exact density functionals and approximations to them. Lieb and Oxford (1981) showed that there was a lower bound to the exchange-correlation energy of the form

$$E_{xc} \geq -C \int dr \frac{n(r)^{4/3}}{r}$$

with $C = 1.68$. The Hellmann-Feynman and virial theorems, and the requirements of coordinate scaling $[n_{\alpha}(r) = \gamma^{-3} n(\gamma r)]$ with $\gamma > 0$, provide tight constraints on approximate forms of $E_{xc}$ and the corresponding potential $v_{xc}$ (Levy and Perdew, 1985; Levy, 1991). Asymptotically exact results for the charge and spin densities and for the effective potential far from finite systems and surfaces were provided by Levy, Perdew, and Sahni (1984) and Almbladh and von Barth (1985).

For decades, chemists have used the energy eigenvalues and the corresponding orbitals to interpret the results of Hartree-Fock calculations for atoms, molecules, and molecular fragments, and energy eigenvalues determined in band structure calculations are essential when condensed matter physicists discuss electronic properties. It is natural to ask whether the Kohn-Sham eigenvalues $\epsilon_i$ in Eq. (14) can be used in similar ways. In the derivation in Sec. IV.A, $\phi_i$ and $\epsilon_i$ are auxiliary variables, but they are more than this.

The identification of the Lagrange parameter $\mu$ in Eq. (11) as the chemical potential means that the highest occupied orbital in exact DF theory should be equal in magnitude to the first ionization energy (Perdew et al., 1982; Perdew and Levy, 1983; Sham and Schlüter, 1983). However, the energy gap between highest occupied and lowest unoccupied orbitals in DF calculations for semiconductors and insulators is usually significantly [about 40%, (Perdew and Levy, 1983)] less than the fundamental gap. This is true for the exact exchange-correlation energy functional as well, since the functional derivative $\delta E_{xc}/\delta n(r)$ has discontinuities as the number of electrons passes through integer values (Perdew and Levy, 1983; Sham and Schlüter, 1983).

One development that had a large impact on my own work was the combination of DF calculations with molecular dynamics introduced by Car and Parrinello (1985). It made DF simulations of bulk systems possible at elevated temperatures, and I give two examples in Sec. VLB. Simulated annealing techniques could also be used to study the energy surfaces of molecules and clusters, where unexpected structures resulted (Jones, 1991). An essential part of DF work prior to 1990 was, of course, the gradual generation of a data base of results for molecules and clusters, as well as extended systems.

3. Functional approximations

Most of the early DF calculations on small clusters and molecules used the LD and/or LSD approximations, often based on the quantum Monte Carlo (QMC) calculations of the correlation energy in the homogeneous electron gas by Ceperley and Alder (1980). Although the results were generally encouraging, it soon became clear that local density calculations can lead to unacceptable errors. Examples were the exchange energy difference between states with different nodal structures (Gunnarsson and Jones, 1985), including the $s-p$ promotion energies in first-row atoms, particularly O and F, $s-d$ promotion energies in transition element atoms, and $d-f$ promotion energies in rare earth atoms. In general, states where the orbitals have the minimum number of nodal planes consistent with the sum rule [Eq. (24)], local density approximations overestimate the interelectronic exchange energy and contribute to the overestimates of binding energies common in small molecules. For excitations requiring the creation of more nodal planes than required by Eq. (24), these

---

24For an introduction to Legendre transforms in this context, see Kutzelnigg (2006).
25A tighter bound on the exchange-correlation energy was given by Chan and Handy (1999). See also Odashima and Capelle (2007).
proximations often greatly underestimate the exchange energy difference, even if the changes in density and spin density are small (Gunnarsson and Jones, 1985).

Dispersion forces—the weak, nonlocal interactions between closed shell systems—pose particular problems for local density approximations. The long-range attraction between separated atoms or molecules is not described, and yet the LD approximation overestimates the binding energy in many such systems, e.g., He₂ (Jones, 1979) and crystalline polyethylene (Montanari and Jones, 1997). However, these approximations give a reasonable description of cohesive properties of crystalline Ar, Kr, and Xe (Trickey, Green, and Averill, 1973).

It is not surprising that better approximations were developed. Density gradients were considered already by Hohenberg and Kohn (1964), and an expression for the correlation energy of an electron gas with high but slowly varying density was derived by Ma and Brueckner (1968). Gradient corrections to the local density exchange approximation were described by Herman, Van Dyke, and Ortenburger (1969). Analysis of the gradient approximation for \( E_{xc} \) in terms of the wave vectors of the contributing fluctuations (Langreth and Perdew, 1977, 1980) showed that the main contribution came from a small region of \( k \) space near the origin. This suggested that the random phase approximation (RPA), which describes this region well, would provide a useful generalization of the gradient approximation.

The sum rules on the exchange-correlation hole discussed above can also be exploited. For example, the spurious oscillations of the second-order gradient expansion of the exchange-hole density \( n_{\text{xc}}(\mathbf{r}, \mathbf{R}) \) for large \( \mathbf{R} \) could be cut off in real space to satisfy the sum rules (Perdew, 1985). Corrections involving density gradients were developed for the correlation (Perdew, 1986; Lee, Yang, and Parr, 1988) and exchange energies (Perdew and Wang, 1986; Becke, 1988). The semi-empirical approximation of Becke was constructed so that the canonical exchange energy density had the correct asymptotic behavior \((\sim r^{-1})\) for atoms.²⁷

In the Hartree-Fock approximation, there is no interaction of an electron with itself, but this self-interaction is canceled imperfectly by the LSD approximation to \( E_{xc} \). There is a long history of work in which this unphysical interaction is canceled explicitly and the remainder approximated [see Gunnarsson and Jones (1981) and references therein, and Perdew and Zunger (1981)]. Total energies of atoms and ions are significantly better, but energy differences between valence configurations showed less improvement (Gunnarsson and Jones, 1981). Finally, we note that a simplification of the Kohn-Sham scheme for calculating the total energy of weakly interacting fragments was developed by Harris (1985). The density is approximated by that of the overlapped fragments, and the absence of spin and the need for self-consistency resulted in a rapid scheme that has been useful in numerous contexts.

It may be difficult for newcomers to the field today to realize just how important these and related developments were to the continuing progress in and ultimate acceptance of DF methods. Nevertheless, the actual number of publications identifying themselves as “density functional” remained small.

### B. Condensed matter

Solid state electronic structure theorists were generally pleased to have justification for the local density calculations they had been performing for years, and a good many moved from electronic band structure, Xα, or Hartree-Fock-Slater calculations into the density functional world. Nevertheless, the mainstream of condensed matter theory was focused elsewhere, for example, on superconductivity and other phase transitions. Figure 1 shows that there the number of DF publications was surprisingly small up to 1990.²⁸ Heine (2002), a prominent condensed matter theorist, looked back on the 1960s in this way:

> “Of course at the beginning of the 1960s the big event was the Kohn Hohenberg Sham reformulation of quantum mechanics in terms of density functional theory (DFT). Well, we recognize it now as a big event, but it did not seem so at the time. That was the second big mistake of my life, not to see its importance, but then neither did the authors judging from the talks they gave, nor anyone else. Did you ever wonder why they never did any calculations with it?”²⁹

There were also prominent critics of density functional and related computational techniques, and one of the best known solid state theoreticians, Anderson (1980) commented:

> “There is a school which essentially accepts the idea that nothing further is to be learned in terms of genuine fundamentals and all that is left for us to do is calculate .... One is left, in order to explain any phenomenon occurring in ordinary matter, only with the problem of doing sufficiently accurate calculations. This is then the idea that I call ‘The Great Solid State Physics Dream Machine’ .... This attitude is closely associated with work in a second field called quantum chemistry.”

This article “never found a publisher” in the U.S. (Anderson, 2011, p. 126), and the only version in print is a French translation (Anderson, 1980). Anderson associated the “Dream Machine” with the name of John Slater (Anderson, 2011, pp. 125 and 126) and described the DF method as a “simplified rather mechanical kind of apparatus” that “shows disturbing signs of becoming a victim of the ‘Dream Machine’ syndrome” (Anderson, 1980). While noting that DF calculations can be particularly valuable in some contexts, he continued:

>²⁷ Although the asymptotic form of the corresponding potential \((\sim r^{-2})\) did not (van Leeuwen and Baerends, 1994).

²⁸I have noted that there were fewer than 90 as late as 1985, when DFT meant “discrete Fourier transform.”

²⁹This is an underestimate. See, for example, Tong and Sham (1966) (atoms) and Lang and Kohn (1970) (jellium surfaces).
“... a great deal of the physics is concealed inside the machinery of the technique, and that very often once one has the answers that these techniques provide, one is not exactly clear what the source of these answers is. In other words the better the machinery, the more likely it is to conceal the workings of nature, in the sense that it simply gives you the experimental answer without telling you why the experimental answer is true.”

These comments seem a little harsh,30 but the increasing availability of computing resources allowed calculations that had previously been impossible, and not all users of the method were critical of the approximations involved.

C. Chemistry

It took many years for DF calculations to be taken seriously by most chemists,31 and the reasons were often convincing: (1) Unlike the TF theory, the Kohn-Sham expression for the energy is not an explicit functional of the density, since the kinetic energy term is defined by an effective potential that leads to the density via a set of orbitals, (2) the original functional of Hohenberg and Kohn is not even defined for all $n$, because not all densities can be derived from the ground state of some single-particle potential (Levy, 1982; Lieb, 1983), and (3) approximations to the exchange-correlation energy are unavoidable, and their usefulness can be assessed only by trying them out. Approximations based on calculations for a homogeneous electron gas, in particular, were unfamiliar to users of atomic orbitals and the Hartree-Fock approximation.32 (4) There was no “systematic” (i.e., mechanical) procedure for refining the results to give the exact solution of the Schrödinger equation and, of course, the exact energy.

This last point was emphasized by many. In principle, the Hartree-Fock method could be extended to multiple determinants (“configuration interaction”) and, coupled with a large basis set, lead to the exact wave function and all properties obtainable from it. This is an attractive proposition, and the dramatic improvements in computing power (3 orders of magnitude per decade) might make the reservations of Dirac less formidable. It was often emphasized that solutions of the Schrödinger equation led to the “right answer for the right reason.” Nevertheless, obtaining numerically exact total energies from wave function calculations remains a major challenge to this day, and it is not surprising that several groups looked at alternatives.

Hartree-Fock-Slater ($X\alpha$) calculations on small molecules were carried out from the early 1970s. The original motivation of Evert Jan Baerends and collaborators in Amsterdam [Heijser, van Kessel, and Baerends (1976) and references therein] was to find a numerically efficient approximation to Hartree-Fock calculations. However, calculations with $\alpha = 0.7$ generally agreed better with experiment than did Hartree-Fock calculations, provided that the full potential was used, not just the “muffin-tin” component then common in many $X\alpha$ calculations.

Some of the first DF calculations on small molecules were performed by Gunnarson and Johansson (1976), with promising results. John Harris and I were initially skeptical that the local density approximations would give useful results for molecules in general, but we developed a full-potential linearized muffin-tin orbital code for small molecules and clusters (Gunnarsson, Harris, and Jones, 1977). These calculations led to good geometries and reasonable binding energies in most cases. In spite of the shortcomings of the local density description of $E_{xc}$, calculations could be performed without adjustable parameters on families of molecules and small clusters that had previously been inaccessible.

We extended this work over the following 5–6 years to over 40 molecular systems, including transition metal dimers (Harris and Jones, 1979b), the group 2 dimers (Jones, 1979, 2012), and the energy surfaces of low-lying states of ozone and SO$_2$ (Jones, 1984b, 1985), H$_2$O, NH$_3$, and CO$_2$ (Jones, 1983; Müller, Jones, and Harris, 1983). The group 2 dimers, particularly Be$_2$, and ozone were cases where the results were qualitatively different from Hartree-Fock theory, but in better agreement with experiment. The comparison of the bond strengths in group 14 dimers C$_2$–Pb$_2$ clearly showed the correlation with the tails of the $s$- and $p$-valence wave functions, and an argument based on the Hellmann-Feynman theorem showed why the C-C bond should be strongest and most versatile, forming single, double, and triple bonds (Harris and Jones, 1979a). The consequences of the unusually compact $p$-valence functions in carbon would not have come as a surprise to solid state physicists with a pseudopotential background [see, for example, Austin and Heine (1966)].

There were parallel studies by other groups, including Baerends and co-workers (Post and Baerends, 1982; Dunlap, Connolly, and Sabin, 1979a, 1979b; Becke, 1982, 1985; Painter and Averill, 1982). This work helped to provide a significant database of molecular results that was essential to convince chemists, in particular, that density functional methods could be valuable.33 It is important to note, however, that much of this work was presented as the results of $X\alpha$, Hartree-Fock-Slater, or local density calculations and was filtered out of the data behind Fig. 1. The results were so encouraging that I wrote in 1984 that the DF approach could be the method for calculating bonding properties (Jones, 1984a).

He referred later to “the oxymoron ‘computational physics’” (Anderson, 1999) and wrote that “... more recently ‘theoretical chemistry’ has become more of a service skill” (Anderson, 2011, pp. 113 and 114).

Walter Kohn summarizes his own experiences, particularly with John Pople, in Kohn and Sherrill (2014). For some of my own, see Jones (2012).

On the other hand, condensed matter physicists were well aware of the poor description of metals provided by the Hartree-Fock approximation, particularly the vanishing density of states at the Fermi energy.

Hoffmann (1977) was the first chemist to encourage our DF work on molecules. He emphasized, however, that we needed more than a few calculations on small molecules if we wanted to convince the many skeptics. Calculations on a variety of systems, particularly organic molecules, would be essential.
Most condensed matter physicists were either not interested or not surprised by these developments in molecular calculations, but theoretical chemists remained skeptical or hostile, or they ignored the development of density functional methods completely. A striking exception was Robert Parr, who wrote (Parr, 1983):

There is an attractive whole alternative theory, the density functional theory. And there are a whole set of attractive alternative computational schemes, of Xα or Kohn-Sham type. Such schemes, at least potentially, offer hope of the same quantitative accuracy as the conventional schemes. And they promise much more as well: a perspicuity, an economy of description, a simplicity of interpretation, and a closeness to the classical ideas of structural chemistry. Much research is needed to bring density functional theory to maturity, but there will result a substantially broadened and more useful quantum chemistry.34

This situation continued throughout the 1980s and into the 1990s. Nevertheless, the number of applications of DF and Xα methods continued to grow. The unexpected structures found in our DF-MD calculations of small phosphorus clusters (Jones and Hohl, 1990) have been confirmed by numerous other calculations since.

D. Menton, 1991

The Seventh International Congress of Quantum Chemistry, held in Menton, France, from 2–5 July 1991, marks for me (and others) a major turning point in the fortunes of DF methods, particularly in chemistry (Fig. 5). Density-related methods were discussed in detail, and communication between their proponents and the skeptics improved. When I asked John Pople why he did not like density functional methods, he replied that the total energies were not good enough. This cannot be denied, but I noted our focus on energy differences, such as binding energies, for which there is no variational principle. All such calculations rely on cancellation of errors.

Becke (1988) was awarded the Medal of the International Academy of Quantum Molecular Science for that year, and he described his development of an approximate exchange functional that promised improvements over local density approximations. In his end-of-conference summary, Pople maintained his distance from density-related methods, but the subsequent comments of Roald Hoffmann were more encouraging. It is certainly no coincidence that Pople and his co-workers tested this approximation on atomization energies of small molecules immediately after the meeting. The first results—energy calculations with the functional of Becke (1988) and Lee, Yang, and Parr (1988) using Hartree-Fock densities—were presented at the Sanibel Symposium in April 1992 (Gill et al., 1992) and marked a shift of the Pople group in favor of density functional methods. More results soon followed and surprised many (Johnson, Gill, and Pople, 1992):

“In summary, these initial results indicate that DFT is a promising means of obtaining quantum mechanical atomization energies; here, the DFT methods B-VWN and B-LYP outperformed correlated ab initio methods, which are computationally more expensive.”

and (Johnson, Gill, and Pople, 1993)

“The density functional vibration frequencies compare favorably with the ab initio results, while for atomization energies two of the DFT methods give excellent agreement with experiment and are clearly superior to all other methods considered.”

The ab initio methods mentioned were Hartree-Fock or self-consistent field (SCF), second-order Möller-Plesset (MP2), and quadratic configuration interaction with single and double substitutions. In addition to the growing body of results on molecules and clusters that were beyond the scope of calculations of correlated wave functions, this change in attitude by one of the most prominent theoretical chemists led to a dramatically new attitude toward the DF method in chemistry.35

34I used precisely this quote at the end of a talk on density functional theory given to theoretical chemists in September 1985, thinking that the similar views of a chemist might add weight to my arguments. The comments afterward made it clear that our efforts to convince chemists still had a long way to go.

35Although one of the authors of these papers wrote an obituary on density functional theory not long afterward (Gill, 2001), noting 1993 as the year of its death.
VI. AFTER ACCEPTANCE (1990—PRESENT)

There were over 160 000 publications on the topics (density functional or DFT) between 1990 and February 2015 (Fig. 1). I leave detailed surveys of this vast and rapidly expanding literature to others and refer again to the very recent issue of the Journal of Chemical Physics (Yang, 2014). The aspects I have chosen should be of general interest, and I give two examples of the possibilities provided by the combination of DF calculations with molecular dynamics.

A. Progress and problems

One of the first signs of growing acceptance of DF methods in chemistry was the incorporation of such calculations into popular ab initio program packages, with Gaussian leading the way. Michael Frisch, the first author of that package, seems to have been a willing convert. At the end of his talk at the ACS National Meeting in San Francisco (13 April 1997) on “ab initio calculations of vibrational circular dichroism and infrared spectra using SCF, MP2, and density functional theories for a series of molecules,” a member of the audience asked:

“What about Hartree-Fock?”

The answer was memorable, and I wrote it down immediately:

“It does not matter what you want to calculate, and it does not matter what functional you use; density functional results are always better than Hartree-Fock.” 36

The availability of such codes and the possibility of comparing the results of different types of calculation were important to establishing the credentials of DF calculations in chemistry.

There has been progress in all the above areas. The number of publications in a year that combine DF calculations with molecular dynamics 37 has grown nearly linearly from almost zero in 1990 to over 2000 in 2014 (Mavropoulos, 2015). Many applications now use Born-Oppenheimer molecular dynamics (Barnett and Landman, 1993), which allow substantially longer time steps in the simulations. The optimized dynamics (Barnett and Landman, 1993), which allow sub-

2001), but DF theory came to be viewed by some as “semi-empirical” in nature. The additional parameters led to improvement over the LD and LSD results, and the use of “training sets” of atomic and molecular systems to optimize the parameters improved the calculated results for particular sets of molecules (Zhao and Truhlar, 2008).

The increased focus on Hartree-Fock–like exchange and its role in hybrid functionals has been one of the most significant developments in recent years. The use of the standard technique of chemists has brought the world of chemistry and materials science closer, as the need for approximations that gave satisfactory results in both areas became obvious. 39

An example of recent studies of functionals that give broad agreement in both fields is that of Peverati and Truhlar (2012). The implementation of HF-like exchange usually comes with

36This “rule” has exceptions, of course.
37Web of Science Core Collection: “Car-Parrinello” or “ab initio first principles MD.”

[the quantum mechanics–molecular mechanics (QM-MM) approach, Carloni, Röthlisberger, and Parrinello (2002)] has found applications in many systems in biology, organic, and solid state chemistry (Lin and Truhlar, 2007; Brunk and Rothlisberger, 2015). Classical force fields that lead to simulations with near-DF accuracy can be developed by a neural network representation of the results of (extremely many) DF calculations on small systems (Behler and Parrinello, 2007), and Li, Kermode, and De Vita (2015) presented recently a scheme where machine learning of forces is used “on the fly” to accelerate DF-MD calculations. Machine learning has also been tested in the context of determining approximate density functionals (Snyder et al., 2012). These and other developments are naturally very welcome.

1. Approximations for $E_{xc}$

The LD [Eq. (17)] and LSD [Eq. (18)] approximations lead to overbinding of many molecules, poor exchange energy differences if the nodal structures of the orbitals change, and the corresponding Kohn-Sham eigenvalues often underestimate measured optical band gaps. Nevertheless, calculations that used them provided insight into many physical problems, and the reasons for the errors (and ways to assess their magnitude) became clearer. However, if insight is not enough and reliable numbers are needed, improved approximations are essential. They have proliferated in recent years.

The first generalized gradient approximations (Perdew, 1986; Becke, 1988; Lee, Yang, and Parr, 1988) did lead to better results, and “hybrid” functionals, which include a Hartree-Fock–like exchange component, were introduced by Becke (1993). 38 This form of $E_x$ has three parameters, and its combination with $E_c$ of Lee, Yang, and Parr (1988) (B3LYP) is still the most common approximation used in chemical applications (Burke, 2012). Many other empirical and hybrid functionals have been developed since, with parameters often fit to thermochemical data for particular groups of molecules. The use of experimental data for fitting functional forms may be understandable (Boese and Handy, 2001), but DF theory came to be viewed by some as “semi-empirical” in nature. The additional parameters led to improvement over the LD and LSD results, and the use of “training sets” of atomic and molecular systems to optimize the parameters improved the calculated results for particular sets of molecules (Zhao and Truhlar, 2008).

The implementation of HF-like exchange usually comes with
a high computational price, especially in calculations using large plane-wave basis sets. However, new algorithms on massively parallel computers have reduced or even eliminated this drawback in some cases (Curioni, 2013).

An alternative path has been followed by others, particularly Perdew and collaborators, who developed a sequence (“Jacob’s ladder”) of approximations without experimental input, where each “rung” builds on the experience of lower levels and satisfies particular physical constraints. The generalized gradient approximation (GGA) of Perdew, Burke, and Ernzerhof (1996) (PBE) incorporates the LSD form below it, and the “meta-GGA” form of Tao et al. (2003) (TPSS), where \( n_1 \) and \( n_1 \) are joined by their gradients and the kinetic energy density of the occupied Kohn-Sham orbitals, builds on both. The agreement with experiment improves (and the computational demands increase) as one climbs the “ladder” (Furche and Perdew, 2006), although the physical interpretation may be less clear. Models of the exchange-correlation hole continue to provide a way of developing DF approximations (Giesbertz, van Leeuwen, and von Barth, 2013), and systematic ways of correcting for the results of DF calculations are still being pursued. A recent example is the use of multideterminant wave functions for hybrid functionals (Savin, 2014).

We noted (Sec. V.A.2) that the functional derivative of the exact exchange-correlation energy functional has discontinuities as the number of electrons passes through integer values. Yang and co-workers focused on this property as a way to study approximate functionals and to find improvements. This has proved to be valuable in the context of many-electron self-interaction (Mori-Sánchez, Cohen, and Yang, 2006) and band gap errors (Cohen, Mori-Sánchez, and Yang, 2008a). It is particularly interesting that errors in DF theory show up in the simplest systems, such as atoms and molecules (Cohen, Mori-Sánchez, and Yang, 2008b).

2. Dispersion interactions

Two areas are particular challenges for DF calculations, the first of which is weak dispersion or van der Waals forces. Crystalline polyethylene, where parallel chains of covalently bonded molecules are bound by such forces (Fig. 6), provides an example of the problems that arise.

The structures obtained by complete relaxation of all atomic coordinates are representative for many similar systems. First, the C-C and C-H bond lengths in the chains are reproduced well by the LDA, PBE (Perdew, Burke, and Ernzerhof, 1996), and Becke-Perdew (BP) (Perdew, 1986; Becke, 1988) approximations (Montanari, Ballone, and Jones, 1998). The C-H bonds are slightly longer than the experimental values for orthorhombic hexatriacontane \( C_{36}H_{74} \) (Teare, 1959), while the C-C bond lengths are reproduced within the experimental uncertainty. The LDA calculations give a bond length \( \sim 1\% \) shorter than in the PBE and BP calculations, and the C-C bond angle is farthest from the measured value. The shortest interchain C-C distance, however, depends dramatically on the choice of functional. The value for the LD approximation (3.67 Å) is much less than the experimental value (4.15 Å), the PBE value (4.66 Å) is too large, and the Becke-Perdew functional results in no binding at all.

The poor results obtained for van der Waals bonds with “standard” approximations for \( E_{xc} \) led to a pronounced increase in dispersion-corrected DF studies in recent years. The progress that has resulted (Dobson and Gould, 2012; Klimeš and Michaelides, 2012; DiStasio, Gobre, and Tkatchenko, 2014) is seen by many as one of the success stories in recent DF research.

The development of a functional that changes seamlessly on going from weakly interacting units to a combined system has been a common goal, and van der Waals interactions can be built into the electron gas picture in various ways (Rapcewicz and Ashcroft, 1991; Dobson and Dinte, 1996; Berland et al., 2014). We have seen (Sec. IV.C) that the exchange and correlation energy can be expressed as an integral over the coupling constant \( \lambda = e^2 \), and this is the basis of several approximations for dispersion (Dobson and Wang, 1999; Dion et al., 2004; Vydrov and Van Voorhis, 2009). The functional of Langreth and co-workers (Dion et al., 2004) is free of experimental input, involving results from electron gas slabs and the electron gas itself, and it has been implemented in several program packages. An empirical correction to DF results (DFT-D2) was proposed by Grimme (2006), and an
alternative was suggested by Tkatchenko and Scheffler (2009). A more recent parametrization by Grimme and co-workers (DFT-D3) (Grimme et al., 2010) provides a consistent description of dispersion forces for elements up to plutonium that can be added to DF packages. It is sensitive to the environment of the atoms involved and has been applied in a range of contexts.

The failure of local density approximations to describe dispersion forces has led to numerous studies of the reasons. One approach has involved the accurate determination of the wave function and correlation effects in simple systems, e.g., He$_2$ (Allen and Tozer, 2002), another uses adiabatic connection [Tkatchenko, Ambrosetti, and DiStasio (2013) and references therein]. Adiabatic connection is also the basis of a range separation approach, where short-range density functional and the long-range random phase approximations are combined (Toulouse et al., 2009). This method describes the bonds in Be and Ne dimers well.

3. Excitations and eigenvalues

The main focus of this article is on DF theory with occupied orbitals, the “Kohn-Sham” aspect. Time-dependent DF theory, based on the theorem of Runge and Gross (1984), has proved to be valuable in predicting the excited states and optical properties in molecules (Burke, Werschnik, and Gross, 2005; Marques et al., 2006; Botti et al., 2007; Isegawa, Peverati, and Truhlar, 2012) and has become an option in most DF program packages. In addition, the exact KS potential can be determined if the exact wave functions are known, and Gritsenko, Schipper, and Baerends (1997) showed that eigenvalue differences provide an excellent approximation to excitation energies in several small molecules when calculated with this potential.

An approximate Kohn-Sham exchange-correlation potential was developed by Schipper et al. (2000) and used within time-dependent DFT to calculate excitation energies and other optical properties of small molecules. The results are in good agreement with experiment, and the model potential appears to approximate the exact KS potential better than the LD, GGA, or meta-GGA forms. Recently, van Meer, Gritsenko, and Baerends (2014) showed for the model potential that the occupied orbital energies are close to ionization energies and virtual-occupied orbital energy gaps are close to excitation energies, leading to a simple physical interpretation of most excitations as single orbital transitions.

4. “Strongly correlated” systems

The term “strongly correlated” is often used in extended systems as a synonym for cases where standard approximations in DF calculations give poor answers. They are usually cases where the potential energy dominates over the kinetic energy and often involve transition element or rare earth atoms. The metal-insulator transition of Mott (Mott, 1968; Imada, Fujimori, and Tokura, 1998) is a well-studied example. Local density approximations (including semilocal approximations, such as GGA) can give qualitatively incorrect descriptions of these materials, examples being metallic DF band structures in insulating transition metal oxides.

The use of model Hamiltonians has been a popular way to treat strong correlations, and a common way of modifying DF calculations is the addition of an on-site Coulomb repulsion (“Hubbard $U$”) in the “LSD + $U$” scheme (Anisimov, Zaanen, and Andersen, 1991; Anisimov, Aryasetiawan, and Lichtenstein, 1997). The parameter $U$ can be estimated within a DF framework (Cococcioni and de Gironcoli, 2005), but it is often fit to experiment. The addition of such Hartree-Fock–like local correlations has a similar effect to hybrid energy functionals (Sec. VI.A.1), where part of the LD exchange energy is replaced by the nonlocal exchange energy. Aras and Kılıç (2014) proposed recently that a range-separated hybrid functional (Heyd, Scuseria, and Ernzerhof, 2003) be combined with a value of $U$ adjusted to reproduce the experimental band gap, and they presented results for some metal chalcogenides. The dominance of the electron-electron interaction over the kinetic energy suggests performing a coupling-constant integration over the latter, and this was formulated by Gori-Giorgi, Seidl, and Vignale (2009).

The dynamical mean-field theory (DMFT) (Georges et al., 1996; Kotliar and Vollhardt, 2004) is a mapping of lattice models onto quantum impurity models subject to a self-consistency condition, and the “LDA + DMFT” method merges the DF theory with a modern many-body approach. It allows the modeling of structural, electronic, and magnetic properties of transition metals, their oxides, and other strongly correlated materials.

In the chemical context, strong correlations are present in systems that cannot be described well by single-determinant solutions to the Kohn-Sham equations. The basic DF theorems tell us that a single symmetry-restricted Slater determinant should be able, in principle, to describe even molecular dissociation, which is a well-known deficiency of Hartree-Fock theory, but the development of appropriate functionals remains a challenge. The study of two-determinant mixing with a recently proposed functional is promising (Becke, 2013).

The adiabatic connection between noninteracting and interacting systems was used [Eqs. (21) and (22)] to show that local density approximations can give reasonable results for systems that are far from uniform, but it can also be used to understand why local density approximations sometimes fail badly. The integrand leading to $E_{xc}$ must, of course, be well behaved mathematically for all $\lambda$. If it depends approximately linearly on $\lambda$, the “mapping” of the noninteracting to the interacting system is straightforward, and standard DF prescriptions should provide a reasonable description. The $\lambda$ dependence of the adiabatic coupling integrand has been studied in the context of hybrid functionals [see, for example, Ernzerhof (1996)] and in calculations for atoms and molecules with up to ten atoms using accurate wave functions and DF approximations (Teale, Coriani, and Helgaker, 2010). If the nature of the system changes for small values of $\lambda$, the linear dependence is a poor approximation, and standard methods fail. Modeling of the integrand to reproduce the correlation energies found in accurate wave function calculations is a promising method of developing new functional approximations.

5. Developments related to QMC

There are developments in the QMC studies of interacting electron systems that could be relevant for future DF work.
The full configuration interaction implementation of QMC was applied recently to the homogeneous electron gas (Shepherd, Booth, and Alavi, 2012) and to simple solids (Booth et al., 2013). Condensed matter scientists have much experience with periodic boundary conditions and plane-wave orbital expansions, and this should aid the implementation of the method in extended systems. Another example is the reformulation of the constrained search approach in DF theory (Levy, 1979; Lieb, 1983) in terms of the density and the \((N - 1)\)-conditional probability density, which can be treated by ground state path integral QMC (Delle Site, Giringhelli, and Ceperley, 2013). It remains to be seen whether the computational demands usually associated with QMC can be reduced.

6. Reduced density matrix functional theory

An active area of research in many-electron systems is based on the one-particle reduced density matrix (1-RDM)

\[
\gamma(r, r') = N \int d\mathbf{r}_2 \cdots d\mathbf{r}_N \Psi(r, \mathbf{r}_2, \ldots, \mathbf{r}_N) \Psi(r', \mathbf{r}_2, \ldots, \mathbf{r}_N),
\]

(29)

where \(\Psi(r_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)\) is the wave function of the \(N\)-electron system and the spin coordinates are suppressed. Gilbert (1975) extended the Hohenberg-Kohn theorem to cover external nonlocal potentials and showed that a nondegenerate ground state wave function \(\Psi_{GR}\) is a universal functional of \(\gamma(r, r')\). The total energy \(E\) can be expressed as a functional of \(\gamma\) whose minimum yields the exact \(\gamma\) and the exact ground state energy for a given external potential [see also Levy (1979)]. In some sense, Gilbert and Levy proved in the modern DF context the idea of Dirac (1930a) to use the 1-RDM to avoid calculating the many-electron wave function. The use of the density matrix provides a description of static correlation by fractional occupation numbers for the orbitals and more flexibility than DF theory. In practice, the exchange-correlation energy is an unknown functional of \(\gamma\), and approximations are unavoidable (Müller, 1984; Buijse and Baerends, 2002). There have been impressive results for geometries of small molecules (Gritsenko, Pernal, and Baerends, 2005) and the energy gaps in bulk semiconductors and transition metal oxides (Sharma et al., 2008).

The obvious similarities between the DF and 1-RDM approaches cannot hide two pronounced differences: The kinetic energy functional in the latter is the exact kinetic energy, which results in poor basis set convergence. Unlike the noninteracting \(\gamma(r, r')\), the interacting analog is also not idempotent, \(^40\) so that there is no artificial (Kohn-Sham) system that reproduces the exact \(\gamma(r, r')\). This may disturb the advocates of density functional calculations, but it may be possible to combine the two approaches and have the advantages of both (Blöchl, Walther, and Pruschke, 2011). In particular, separation of the short-range density functional from the long-range density matrix functional is a promising development (Rohr, Toulouse, and Pernal, 2010).

7. “Missing” topics

A review of this length cannot cover all density functional topics, even those of current interest and importance. My main focus has been on the “occupied orbitals only” world of Kohn and Sham (1965) and calculations of the total energy, but this has not deterred readers of earlier versions from drawing my attention to omissions. There have been many improvements in numerical methods and algorithms, particularly those that scale linearly with the number of particles \(N\). Examples are the kKRnano code (Zeller, 2008), which has been used to simulate many thousands of particles, and orbital-free methods [see Karasiev, Sjostrom, and Trickey (2014) and references therein]. The development of new functional approximations is essential in the latter case. A glance at the program of any conference related to density functional methods immediately shows a large range of applications of these methods. I encourage interested readers to look again at the recent volume (Yang, 2014).

8. Terminology

The terms \textit{ab initio} and “first principles” are used differently in the “chemical” and “materials” worlds. For most chemists, these expressions mean solutions of the Schrödinger equation for the system of interacting electrons (e.g., by QMC), for many materials scientists it can be a DF calculation without (or even with) adjustable parameters. I use the term “density functional” to describe the work I do, and \textit{ab initio} for solutions of the Schrödinger equation. I have been told countless times over the years that the latter give “the right answer for the right reason,” and I look forward to the day when the speakers say the same about DF calculations.

B. Two applications

I briefly describe the results of one DF simulation in biology and one in materials science that indicate the scale of DF calculations that are possible. They are not standard applications by any means, requiring very large computing resources and substantial human effort. In the spirit of addressing newcomers to the field, I now mention some practical points about DF calculations.

The ready availability of DF programs does not mean that they are easy to use or that careful choices of input parameters can be avoided. Some methods of calculation consider all electrons, but many focus on the outermost valence electrons that dominate structural properties and replace the effect of inner “core” electrons by an effective potential (pseudopotential) or a frozen core density. These must be constructed as needed and tested. A basis set must be chosen and its convergence checked, and the final results depend, in particular, on the approximation adopted for the exchange-correlation energy functional. This means that numerous test calculations are essential before the actual calculations of interest can be carried out. The following applications use the PBE functional (Perdew, Burke, and Ernzerhof, 1996), which is used widely in the materials world.

In 1871, Charles Darwin saw the formation of protein molecules under extreme, prebiotic, conditions as a possible path of evolution of life on Earth:
“But if (and what a big if) we could conceive in some warm little pond with all sorts of ammonia and phosphoric salts—light, heat, electricity, &c present, that a protein compound was chemically formed, ready to undergo still more complex changes, at the present day such matter would be instantly devoured, or absorbed, which would not have been the case before living creatures were formed” (Darwin, 1887).

DF simulations can be used to test whether biological molecules can be produced by chemical processes that do not require biological synthesis machines like ribosomes. Schreiner et al. (Nair, Schreiner, and Marx, 2008; Schreiner, Nair, and Marx, 2008; Schreiner et al., 2011) studied possible reactions of N-carboxy anhydrides (a form of activated amino acids) in water under high pressures and temperatures in the presence of pyrites FeS$_2$ [the controversial “iron-sulfur world” of Wächtershäuser (1988)]. The presence of an FeS$_2$ surface changes the free energetics of the steps of the carbonyl sulfide (COS)-mediated polymerization of glycine carried out under different thermodynamic conditions (Fig. 7), and it stabilizes the peptide product against hydrolysis.

The reactions studied are just a few of many possible scenarios for the production of molecules that are essential to life on Earth, but they demonstrate the value of simulations under conditions that are difficult to attain experimentally. They also show that simulations without adjustable parameters can be performed on biological systems that were unthinkable with earlier generations of computers. Of course, there are many such systems for which the simulation sample sizes currently accessible with DF methods are simply inadequate. Classical force fields with appropriately chosen parameters are likely to remain the method of choice for such systems for some time yet.

Phase change (PC) materials are alloys of chalcogens (group 16 elements) that are ubiquitous in the world of rewritable optical storage media, examples being the digital versatile disk (DVD-RW) and Blu-ray Disc. Nanosized bits in a thin polycrystalline layer are switched reversibly and extremely rapidly between amorphous and crystalline states, and the state can be identified by changes in resistivity or optical properties. Crystallization of the amorphous bit is the rate-limiting step in the write or erase cycle, and much attention has been focused on this process. Alloys of Ge, Sb, and Te are often used in PC materials, and 460-atom simulations have been carried out at 600 K on amorphous Ge$_2$Sb$_2$Te$_5$ (GST) (Kalikka et al., 2012; Kalikka, Akola, and Jones, 2014) (Fig. 8). Crystallization takes place in just over 1 ns, and it is possible to monitor changes in the distribution of the cavities, the diffusion of atoms of the different elements, and percolation of crystalline units in the sample. The presence of cavities is essential for rapid crystallization in these materials (Akola and Jones, 2007, 2008).

These calculations required over 400,000 self-consistent DF calculations of energies and forces for a 460-atom sample, and it was exciting to witness the gradual transition from an amorphous sample to a crystal without driving it in any way. It is very unusual for a phase transition to be fast enough (nanoseconds in this case) to be accessible to DF calculations under the actual conditions of temperature and time scale, and the details of the structural change should shed light on the mechanisms of other phase transitions. I emphasize again that the steady improvement in numerical algorithms has played an important role in making such calculations possible, but they also require computers of the highest performance class. Furthermore, these and almost all other simulations use a functional approximation developed for $T = 0$ K. A recent parametrization of restricted path integral Monte Carlo data for the homogeneous electron gas for wide temperature and density ranges (Karasiev et al., 2014) indicates that improved
functionals for describing systems at high temperatures should soon be available.

C. DF theory in other contexts

The dominant role of the electron density functional theory in materials science and chemistry means that many of its practitioners identify the term solely with electrons. “DF calculations” are also performed in other contexts.

1. Classical density functional theory

Work in this area focuses on the thermodynamics of systems interacting with classical potentials and is based on the Hohenberg-Kohn-Mermin formulation at finite temperatures (Sec. V.A). The focus is on the Helmholtz free energy and its expression as a functional of the density. The systems considered are very large by electron density functional standards (I have read about particle numbers “up to Avogadro’s number”), with particular emphasis on phase transitions. Early work was performed by Nordholm and co-workers (Nordholm and Haymet, 1980; Johnson and Nordholm, 1981), who described a “generalized van der Waals theory” using a free energy functional defined in terms of a coarse-grained particle density. Surveys of the field show that the classical and electronic DF worlds have much in common (Evans, 1979, 1992; Haymet and Oxtoby, 1981; Baus, 1990; Ashcroft, 1995; Wu and Li, 2007). Electron density functional theory needs approximations for the exchange-correlation energy, the classical counterpart an expression for the excess free energy arising from interactions in the system. The free energy is written as the sum of terms involving a reference system (e.g., a liquid) and a second system of interest (such as a solid), and the “coupling-constant integration” technique can be used to switch from one density to the other or from a “noninteracting” to an “interacting” system.

Particularly familiar to me were the weighted density (WD) approximations for the excess free energy in an inhomogeneous liquid (Tarazona, 1984; Curtin and Ashcroft, 1985), where the free energy density at a given point is taken to be that of a homogeneous system with a density determined by weighting over a physically relevant region. This is precisely the motivation of the “WD approximation” in electronic systems (Gunnarsson, Jonson, and Lundqvist, 1979; Gunnarsson and Jones, 1980), where the “physically relevant region” is determined by the exchange-correlation hole [Eq. (22)]. In the WD approximation, this can be written

\[ n_{\text{sc}}^{\text{WD}}(\mathbf{r}, \mathbf{r}' - \mathbf{r}) \equiv n(\mathbf{r}')G^{\text{WD}}(\mathbf{r}, \mathbf{r}'; \tilde{n}(\mathbf{r})). \]  

where \( G^{\text{WD}} \) is a model pair-correlation function, and \( \tilde{n}(\mathbf{r}) \) is a nonlocal parameter chosen to satisfy the sum rule Eq. (24). \( G^{\text{WD}} \) can be chosen to satisfy additional requirements, such as reproducing the LD result in a homogeneous system or the correct form of the potential outside a surface. A particularly simple form has been tested for atoms with mixed success (Gunnarsson and Jones, 1980; Jones and Gunnarsson, 1989).

2. Nuclei

The number of nuclides (isotopes) of the elements that occur or can be synthesized on Earth is approximately 3000 and increases every year. Of these, less than 300 are considered stable (Erler et al., 2012). The stability of a nuclide is determined by the amount of energy needed to remove a single neutron or proton or a pair of neutrons or protons, and it is natural that the limits of nuclear binding have received much attention. The microscopic method of choice in describing weakly bound complex nuclei is the nuclear density functional theory based on a self-consistent mean-field approach (Bender, Heenen, and Reinhard, 2003).

There has been significant progress in developing a density functional theory for schematic nucleon-nucleon interactions along the lines discussed above for electrons (Drut, Furnstahl, and Platter, 2010; Drut and Platter, 2011). Nevertheless, under many-body theories are not yet able to provide input for effective energy-density functionals, and experimental input cannot be avoided. Many calculations are based on the functional form suggested by Skyrme (1956, 1959), and the optimization of the parameters involved is an ongoing challenge (Klüpfel et al., 2009; Kortelainen et al., 2010; Carlsson, Toivanen, and von Barth, 2013). A detailed study of the “nuclear landscape” using a range of functionals concluded that the number of bound nuclides with between 2 and 120 protons is \( 6900 \pm 500 \) (Erler et al., 2012).

Numerous concepts that one encounters in the nuclear physics literature [such as Skyrme-Hartree-Fock, mean-field theory, DFT, Bardeen-Cooper-Schrieffer (BCS) wave function, superfluidity, spin-orbit coupling, orbital-dependent functionals, optimized effective potentials, even Thomas-Fermi] are familiar to many in the condensed matter world. It is indeed interesting to look beyond familiar pastures, and perhaps the experience of condensed matter physicists and chemists with density functional theories can provide important insights in other areas.

VII. SUMMARY AND OUTLOOK

A. An “approximate practical method”

The astonishing growth of density functional calculations since 1990 (Fig. 1) is due in part to the use of this term for calculations that had previously used the labels electron band structure, local density, and Hartree-Fock-Slater.\(^4\) Such calculations would certainly have continued with a variety of local density approximations, but the crucial theoretical justification of the density functional formalism resulted in Walter Kohn sharing the 1998 Nobel Prize for Chemistry with John Pople (Kohn, 1999). Anderson noted that, although “very deep problems” remain, this award may indicate that

the labors and controversies … in understanding the chemical binding in materials had finally come to a resolution in favor of ‘LDA’ and the modern computer” (Anderson, 2011, p. 109).

\(^4\)Some program names were changed correspondingly.
The recognition of the basic role of the electron density goes back to the earliest days of quantum mechanics and the first density functional of Thomas and Fermi. Dirac (1930a) incorporated exchange effects into this model and observed that a matrix density function (we know it as the one-particle reduced density matrix) could completely determine the state of an atom within the Hartree-Fock approximation; “it is not necessary to specify the individual three-dimensional wave functions that make up the total electric density.” The self-consistent field of Hartree and the use of determinantal wave functions by Slater, Bloch, and Fock were followed by the calculations of Wigner and Seitz, who showed that the exchange hole of an electron in Na metal was localized to a single ion. The exchange-hole picture is a central concept in DF theory and was developed further by Slater in 1951. An effective exchange potential of the form of Bloch and Dirac was derived and tested by Gáspár in 1954. Hohenberg and Kohn (1965) and Kohn and Sham (1965) are two of the most cited papers in physics.

Kohn and Sham (1965) did not expect the local density approximation for exchange and correlation to give “an accurate description of chemical binding,” and this view was shared by most theoretical chemists for many years. Nevertheless, we now know that approximations to the exchange-correlation energy based on the homogeneous electron gas (LD, LSD) and their modifications provide a scheme that can give valuable information about large molecules and extended systems. It was shown later why approximations to $E_{xc}$ could give good results for density distributions far from those where they are obviously valid (Jones and Gunnarsson, 1989). The approach does appear to satisfy the 80-year-old goal of Dirac (1930a) to find “approximate practical methods of applying quantum mechanics to (explain) the main features of complex atomic systems without too much computation.” Those involved in this activity know that few of us took the last point too seriously.

B. Exchange-correlation approximations

The LD and LSD approximations have well-documented drawbacks, and the resulting numbers (binding energies, band gaps, etc.) should always be treated with caution. However, the approximations and their further developments satisfy important physical criteria, including sum rules on the exchange and correlation holes, coordinate scaling properties, and bounds on $E_{xc}$, and our long experience with them helps us to predict when the results may be wrong and by how much. After spending several years developing functional forms based on ever-expanding amounts of experimental data, Handy (2002, 2009) noted that some simpler functionals allow us to separate different types of correlation in molecules and concluded:

“We are returning to the view that the simplest parameter-free GGA functionals are the best functionals to use with DFT, because they offer the simplest interpretation and have greater global predictive power.”

In fact, the bonding patterns are correct in most cases, which is no doubt one reason why LD approximations and their extensions are still in widespread use. They make possible the simultaneous study of numerous related systems, such as families of molecules or materials, with the computational resources needed to determine the wave function of a single much smaller system. Is this pragmatic approach giving way to the search for schemes that produce better numbers automatically, as the long list of approximate functionals and publications comparing their predictions might suggest?

Thirty years ago, it became obvious (to Olle Gunnarsson and me, at least) that the LSD approximation could lead to unacceptably large errors in exchange energy differences if the nodal structures of the orbitals involved changed. Some in the chemical world advised us then to go back to Hartree-Fock calculations and their extensions. Now Becke (2014) writes

“This is my position, today, on how to proceed into the future. We have come ‘full circle’ back to exact exchange or, if you prefer, Hartree-Fock.”

Going around in circles is not usually a sign of progress. Have we overlooked something?

C. Quo vadis?

Figure 1 shows that density functional theory will be with us for the foreseeable future, and we can be sure that some exciting results lie ahead. Why then should two of the most cited workers in the field have serious reservations about the future of DF theory? Burke (2012) wrote that “it is clearly both the best of times and the worst of times for DFT” and wondered whether it is time for a “paradigm shift.” A newcomer to the field might indeed despair of understanding why one of the countless approximations for $E_{xc}$, even those with a sound physical basis, should be favored over another, or the real physical reasons behind a particular result. Are DF calculations now following the “Dream Machine” scenario foreseen many years ago by Anderson (1980)? Furthermore, the identification of the “best” functional may be ambiguous. A comparison of the band gaps in LiH and four alkali halides, four oxides, and solid Ne and Ar (gaps between 0.2 and 20 eV) with the predictions of many popular functionals (Civalleri et al., 2012) showed that finding the best functional also depends on the choice of statistical measure (mean error, mean absolute error, variance, etc.).

The concerns of Becke (2014) are just as real. It is obvious that great progress has been made in applying DF methods to systems that seemed beyond us only 10 or 15 years ago, and the use of Hartree-Fock–like exchange in many modern functionals has helped communication between the different fields where DF methods are used. However, Becke (and many others, including me) have focused for years on the “Kohn-Sham” version of DF theory (“occupied orbitals only”), which is a major reason for the popularity of the method. Should we move our focus away from the relatively featureless electron density with its small and subtle changes? Can the combination with density matrix functional methods lead to a new breakthrough? Will the inherent accuracy of wave-function-based methods prove to be decisive as
computational resources expand?  Many years ago, a colleague predicted that DF methods would ultimately lose out to solutions of the Schrödinger equation as computer power increased. He was not impressed by my view that DF calculations would always be far ahead in the size of system we could calculate (I think I said 5–10 years, but it is more), and he moved on. I stayed.

Density functional theory has a long and fascinating history involving some of the best known names in physics. It deserves better than to be the basis for developing a vast array of approximations seeking the “right” numbers. It cannot be expected to provide precise answers to all questions when simple descriptions of the exchange-correlation energy are used, but its ability to outperform methods that seek exact solutions of the Schrödinger equation is not threatened. We shall continue to obtain insight into all sorts of problems that we cannot imagine today.

ACKNOWLEDGMENTS

Our density functional work in Jülich has involved many collaborators, but the contributions of Olle Gunnarsson and John Harris in the early years were crucial, as were the support of Gert Eilenberger and his conviction that curiosity-driven research was not (yet) an outdated concept. We have benefited greatly over many years from generous grants of time on supercomputers in the Forschungszentrum Jülich by the John von Neumann Institute for Computing, JARA/HPC, and the FZ Jülich. I thank J. Akola and J. Kalikka for collaboration on GST crystallization (Fig. 8), P. Mavropoulos and D. Marx for providing Figs. 1 and 7, respectively, E. J. Baerends, M. Levy, A. Savin, and many others for criticism of earlier versions of the manuscript, and Richard Jones for many discussions and suggestions.

APPENDIX: ADIABATIC COUPLING

If we make infinitesimally slow changes to a parameter describing a system in a stationary state, the system remains in its original state. This adiabatic principle is familiar from the earlier quantum theory (Ehrenfest, 1916) and was shown already in 1926 to be valid in quantum mechanics (Born, 1926b; Born and Fock, 1928). If the Hamiltonian \( H \) of a system depends on some parameter \( \lambda \), e.g., interatomic separation, the electron-electron interaction, or the time, Güttinger (1931) showed that

\[
\frac{\partial E_n}{\partial \lambda} = \left( \frac{\partial H}{\partial \lambda} \right)_{nn} \tag{A1}
\]

for the diagonal terms, where the right-hand side is an expectation value for state \( n \), and

\[
\int dr \ u_n^* \frac{\partial u_n}{\partial \lambda} = \frac{1}{E_n - E_m} \left( \frac{\partial H}{\partial \lambda} \right)_{nn} \tag{A2}
\]

for \( n \neq m \). \( E_n \) is the energy eigenvalue and \( u_n \) is the eigenfunction of state \( n \). These results are implicit in first-order Rayleigh-Schrödinger perturbation theory (Schrödinger, 1926a), and Born and Fock (1928) used off-diagonal terms of Eq. (A2) in their proof of adiabaticity in quantum mechanics. If \( \lambda \) is an interatomic distance in a molecule, Eq. (A1) is familiar as the “Hellmann-Feynman theorem” for forces (Hellmann, 1933, 1937; Feynman, 1939).

The integrated form of Eq. (A1)

\[
E_n(\lambda_f) - E_n(0) = \int_0^{\lambda_f} d\lambda \left( \frac{\partial H}{\partial \lambda} \right)_{nn} \tag{A3}
\]

is often attributed to Pauli [see Musher (1966) and references therein] and described as the “Pauli trick.” However, it is an obvious extension of Eq. (A1) and is an intermediate result of Güttinger in the context of the adiabatic principle, where \( \lambda \) is the time. The “integration over the coupling constant” or “adiabatic coupling” method is an obvious choice in studies of the relationship between the noninteracting (\( \lambda = 0 \)) and interacting (\( \lambda = \epsilon^2 \)) electron systems (see Sec. IV.C).

43The mathematical equivalence of perturbation theory and differentiation was, however, shown only some years later (Rellich, 1937a, 1937b, 1941). See also Pupyshev (2000).

REFERENCES


Curioni, A., 2013 (personal communication).


Hoffmann, R., 1977 (personal communication).


Teare, P. W., 1959, “The crystal structure of orthorhombic hexatriacontane C_{36}H_{78},” Acta Crystallogr. 12, 294–300.


