An Introduction to Density Functional Theory

N. M. Harrison

Department of Chemistry, Imperial College of Science Technology and Medicine, SW7 2AY, London and

CLRC, Daresbury Laboratory, Daresbury, Warrington, WA4 4AD

For the past 30 years density functional theory has been the dominant method for the quantum mechanical simulation of periodic systems. In recent years it has also been adopted by quantum chemists and is now very widely used for the simulation of energy surfaces in molecules. In this lecture we introduce the basic concepts underlying density functional theory and outline the features that have lead to its wide spread adoption. Recent developments in exchange correlation functionals are introduced and the performance of families of functionals reviewed.

The lecture is intended for a researcher with little or no experience of quantum mechanical simulations but with a basic (undergraduate) knowledge of quantum mechanics. We hope to provide sufficient background to enable informed judgements on the applicability of a particular implementation of density functional theory to a specific problem in materials simulation.

For those who wish to go more deeply into the formalism of density functional theory there are a number of reviews and books aimed at intermediate and advanced levels available in the literature [1,2,3]. Where appropriate source articles are referred to in the text.

1. The Solution of the Schrödinger Equation

During the course of this lecture we will be primarily concerned with the calculation of the ground state energy of a collection of atoms. The energy may be computed by solution of the Schrödinger equation – which, in the time independent, non-relativistic, Born-Oppenheimer approximation is[1];

$$\hat{H}\,\Psi(\mathbf{r}_1,\mathbf{r}_2,...,\mathbf{r}_N) = E\Psi(\mathbf{r}_1,\mathbf{r}_2,...,\mathbf{r}_N)$$

Equation 1

[1] Atomic units are used throughout.

The Hamiltonian operator, H, consists of a sum of three terms; the kinetic energy, the interaction with the external potential (V_{ext}) and the electron-electron interaction (V_{ee}). That is;

$$\hat{H} = -\frac{1}{2}\sum_{i}^{N} \nabla_i^{\,2} + \hat{V}_{ext} + \sum_{i<j}^{N} \frac{1}{|\,\mathbf{r}_i - \mathbf{r}_j\,|}$$

Equation 2

In materials simulation the external potential of interest is simply the interaction of the electrons with the atomic nuclei;

$$\hat{V}_{ext} = -\sum_{a}^{N_{at}} \frac{Z_a}{|\,\mathbf{r}_i - \mathbf{R}_a\,|}$$

Equation 3

Here, $\mathbf{r}_i$ is the coordinate of electron *i* and the charge on the nucleus at $\mathbf{R}_a$ is Z_a. Note that in order to simplify the notation and to focus the discussion on the main features of DFT the spin coordinate is omitted here and throughout this article. Equation 1 is solved for a set of Ψ subject to the constraint that that the Ψ are anti-symmetric – they change sign if the coordinates of any two electrons are interchanged. The lowest energy eigenvalue, E_0, is the ground state energy and the probability density of finding an electron with any particular set of coordinates $\{\mathbf{r}_i\}$ is $|\Psi_0|^2$.

The average total energy for a state specified by a particular Ψ, not necessarily one of the eigenfunctions of Equation 1, is the expectation value of H, that is;

$$E[\Psi] = \int \Psi^* \hat{H} \Psi \, d\mathbf{r} \equiv \left\langle \Psi \middle| \hat{H} \middle| \Psi \right\rangle$$

Equation 4

The notation $[\Psi]$ emphasises the fact that the energy is a *functional* of the wavefunction. The energy is higher than that of the ground state unless Ψ corresponds to Ψ_0 – which is the *variational theorem*;

$$E[\Psi] \geq E_0$$

Equation 5

The ground state wavefunction and energy may be found by searching all possible wavefunctions for the one that minimises the total energy. Hartree-Fock theory consists of an ansatz for the structure of Ψ - it is assumed to be an antisymmetric product of functions (f_i) each of which depends in the coordinates of a single electron, that is;

$$\Psi_{HF} = \frac{1}{\sqrt{N!}} \det\left[f_1 f_2 f_3 \ldots f_N \right]$$

Equation 6

where, det indicates a matrix determinant [4]. Substitution of this ansatz for Ψ into the Schrödinger equation results in an expression for the Hartree Fock energy;

$$\begin{aligned} E_{HF} = & \int f_i^*(\mathbf{r}) \left(-\frac{1}{2} \sum_i^N \nabla_i^2 + V_{ext} \right) f_i(\mathbf{r}) d\mathbf{r} \\ & + \frac{1}{2} \sum_{i,j}^N \int \frac{f_i^*(\mathbf{r}_1) f_i(\mathbf{r}_1) f_j^*(\mathbf{r}_2) f_j(\mathbf{r}_2)}{\left| \mathbf{r}_i - \mathbf{r}_j \right|} d\mathbf{r}_1 d\mathbf{r}_2 \\ & - \frac{1}{2} \sum_{i,j}^N \int \frac{f_i^*(\mathbf{r}_1) f_j(\mathbf{r}_1) f_i(\mathbf{r}_2) f_j^*(\mathbf{r}_2)}{\left| \mathbf{r}_i - \mathbf{r}_j \right|} d\mathbf{r}_1 d\mathbf{r}_2 \end{aligned}$$

Equation 7

The second term is simply the classical Coulomb energy written in terms of the orbitals and the third term is the *exchange* energy. The ground state orbitals are determined by applying the variation theorem to this energy expression under the constraint that the orbitals are orthonormal. This leads to the Hartree-Fock (or SCF) equations;

$$\left[-\frac{1}{2} \nabla^2 + v_{ext}(\mathbf{r}) + \int \frac{r(\mathbf{r}')}{\left| \mathbf{r} - \mathbf{r}' \right|} d\mathbf{r}' \right] f_i(\mathbf{r}) + \int v_X(\mathbf{r}, \mathbf{r}') f_i(\mathbf{r}') d\mathbf{r}' = e_i f_i(\mathbf{r})$$

Equation 8

Where the non-local exchange potential, v_X, is such that:

$$\int v_X(\mathbf{r},\mathbf{r}')f_i(\mathbf{r}')d\mathbf{r}' = -\sum_j^N \int \frac{f_j(\mathbf{r})f_j^*(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} f_i(\mathbf{r}')\, d\mathbf{r}'$$

Equation 9

The Hartree-Fock equations describe non-interacting electrons under the influence of a mean field potential consisting of the classical Coulomb potential and a *non-local* exchange potential.

From this starting point better approximations (correlated methods) for Ψ and E_0 are readily obtained but the computational cost of such improvements is very high and scales prohibitively quickly with the number of electrons treated (for an excellent introduction see ref. [4]). In addition, accurate solutions require a very flexible description of the wavefunction's spatial variation, i.e. a large and basis set is required which also adds to the expense for practical calculations. Many correlated methods have been developed for molecular calculations [4]. The cost of the most commonly used methods, MP2, MP3, MP4, CISD, CCSD, CCSD(T) formally scales with the number of electrons raised to the power of 5,6,7,6,6,7 respectively. In most cases CCSD(T) calculations are of sufficient accuracy to determine the chemical properties of systems to sufficient accuracy to predict chemical properties (stability, reaction rates …) however, due to the computational expense the routine application of such methods to realistic models of systems of interest is not practical and not likely to become so despite rapid advances in computer technology. In this context we mention recent advances in the solution of the Schrödinger equation using variational quantum Monte Carlo approach [5].

The discussion above has established that direct solution of the Schrödinger equation is not currently feasible for systems of interest in condensed matter science – this is a major motivation for the development and use of density functional theory.

The question that arises is – *Is it necessary to solve the Schrödinger equation and determine the 3N dimensional wavefunction in order to compute the ground state energy ?*

2. Avoiding the Solution of the Schrödinger Equation

The Hamiltonian operator (Equation 2) consists of single electron and bi-electronic interactions – i.e. operators that involve on the coordinates of one or two electrons only. In order to compute the total energy we do not need to know the 3N dimensional wavefunction. Knowledge of the two-particle probability density – that is, the probability of finding an electron at $\mathbf{r}_1$ and an electron at $\mathbf{r}_2$ is sufficient.

A quantity of great use in analysing the energy expression is the second order density matrix, which is defined as:

$$P_2(\mathbf{r}_1^{'},\mathbf{r}_2^{'};\mathbf{r}_1,\mathbf{r}_2) = \frac{N(N-1)}{2}\int \Psi^*(\mathbf{r}_1^{'},\mathbf{r}_1^{'},...,\mathbf{r}_N^{'})\Psi(\mathbf{r}_1,\mathbf{r}_2,...,\mathbf{r}_N)d\mathbf{r}_3 d\mathbf{r}_4 ... d\mathbf{r}_N$$

Equation 10

The diagonal elements of P_2, often referred to as the *two-particle density matrix* or *pair density*, are;

$$P_2(\mathbf{r}_1,\mathbf{r}_2) = P_2(\mathbf{r}_1,\mathbf{r}_2;\mathbf{r}_1,\mathbf{r}_2)$$

This is the required two electron probability function and completely determines all two particle operators. The first order density matrix is defined in a similar manner and may be written in terms of P_2 as;

$$P_1(\mathbf{r}_1^{'};\mathbf{r}_1) = \frac{2}{N-1}\int P_2(\mathbf{r}_1^{'},\mathbf{r}_2;\mathbf{r}_1,\mathbf{r}_2)d\mathbf{r}_2$$

Equation 11

Given P_1 and P_2 the total energy is determined exactly;

$$E = tr\left(\hat{H}\,\hat{P}\right) = \int\left[\left(-\frac{1}{2}\nabla_1^{\,2} - \sum_a^{N_{at}}\frac{Z_a}{|\mathbf{r}_1-\mathbf{R}_a|}\right)P_1(\mathbf{r}_1^{'},\mathbf{r}_1)\right]_{\mathbf{r}_1=\mathbf{r}_1^{'}} d\mathbf{r}_1 + \int\frac{1}{|\mathbf{r}_1-\mathbf{r}_2|}P_2(\mathbf{r}_1,\mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2$$

Equation 12

We conclude that the diagonal elements of the first and second order density matrices completely determine the total energy. This appears to vastly simplify the task in hand. The solution of the full Schrödinger equation for Ψ is not required – it is sufficient to determine P_1 and P_2 - and the problem in a space of 3N coordinates has been reduced to a problem in a 6 dimensional space.

Approaches based on the direct minimisation of E(P_1,P_2) suffer from the specific problem of ensuring that the density matrices are legal – that is, they must be constructible from an antisymmetric Ψ. Imposing this constraint is non trivial and is currently an unsolved problem [6,7]. In view of this we conclude that Equation 12 does not lead immediately to a reliable method for computing the total energy without calculating the many body wavefunction.

The observation which underpins density functional theory is that we do not even require P_2 to find E – *the ground state energy is completely determined by the diagonal elements of the first order density matrix – the charge density.*

3. The Hohenburg-Kohn Theorems

In 1964 Hohenburg and Kohn proved the two theorems [8]. The first theorem may be stated as follows;

The electron density determines the external potential (to within an additive constant).

If this statement is true then it immediately follows that the *electron density uniquely determines the Hamiltonian operator* (Equation 2). This follows as the Hamiltonian is specified by the external potential and the total number of electrons, N, which can be computed from the density simply by integration over all space. Thus, in principle, given the charge density, the Hamiltonian operator could be uniquely determined and this the wave functions Ψ (of all states) and all material properties computed.

Hohenburg and Kohn [8] gave a straightforward proof of this theorem, which was generalised to include systems with degenerate states in proof given by Levy in 1979 [9]. It is said that the theoretical spectroscopist E. B. Wilson put forward a very straightforward proof of this theorem during a meeting in 1965 at which it was being introduced. Wilson's observation is that the electron density uniquely determines the positions and charges of the nuclei and thus trivially determines the Hamiltonian. This proof is both transparent and elegant – it is based on the fact that the electron density has a cusp at the nucleus, such that;

$$Z_a = \frac{-1}{2\bar{r}(0)}\left[\frac{\partial \bar{r}(r_a)}{\partial r_a}\right]_{r_a=0}$$

where $\bar{r}(\mathbf{r})$ is the spherical average of ρ and so a sufficiently careful examination of the charge density uniquely determines the external potential and thus the Hamiltonian.

Although less general than the Levy proof this observation establishes the theorem for the case of interest – electrons interacting with nuclei. The first theorem may be summarised by saying that the energy is a functional of the density – E[ρ].

The second theorem establishes a variational principle;

For any positive definite trial density, r_t, such that $\int r_t(\mathbf{r})d\mathbf{r} = N$ *then* $E[r_t] \geq E_0$

The proof of this theorem is straightforward. From the first theorem we know that the trial density determines a unique trial Hamiltonian (H_t) and thus wavefunction (Ψ_t); $E[r_t] = \langle \Psi_t \mid H \mid \Psi_t \rangle \geq E_0$ follows immediately from the variational theorem of the Schrödinger equation (Equation 5). This theorem restricts density functional theory to studies of the ground state. A slight extension allows variation to excited

states that can be guaranteed orthogonal to the ground state but in order to achieve this knowledge of the exact ground state wavefunction is required.

The two theorems lead to the fundamental statement of density functional theory;

$$d\left[E[r]-m(\int r(\mathbf{r})d\mathbf{r}-N)\right]=0$$

Equation 13

The ground state energy and density correspond to the minimum of some functional $E[\rho]$ subject to the constraint that the density contains the correct number of electrons. The Lagrange multiplier of this constraint is the electronic *chemical potential* μ.

The above discussion establishes the remarkable fact that there is a *universal* functional $E[\rho]$ (i.e. it does not depend on the external potential which represents the particular system of interest) which, if we knew its form, could be inserted into the above equation and minimised to obtain the *exact* ground state density and energy.

4. The Energy Functional

From the form of the Schrödinger equation (Equation 1) we can see that the energy functional contains three terms – the kinetic energy, the interaction with the external potential and the electron-electron interaction and so we may write the functional as;

$$E[r]=T[r]+V_{ext}[r]+V_{ee}[r]$$

The interaction with the external potential is trivial;

$$V_{ext}[r]=\int \hat{V}_{ext}\, r(\mathbf{r})d\mathbf{r}$$

The kinetic and electron-electron functionals are unknown. If good approximations to these functionals could be found direct minimisation of the energy would be possible; this possibility is the subject of much current research - see for instance Ref. [10].

Kohn and Sham proposed the following approach to approximating the kinetic and electron-electron functionals [11]. They introduced a fictitious system of N non-interacting electrons to be described by a single determinant wavefunction in N

"orbitals" f_i. In this system the kinetic energy and electron density are known exactly from the orbitals;

$$T_s[r] = -\frac{1}{2}\sum_i^N \left\langle f_i \left| \nabla^2 \right| f_i \right\rangle$$

Here the suffix emphasises that this is not the true kinetic energy but is that of a system of non-interacting electrons, which reproduce the true ground state density;

$$r(\mathbf{r}) = \sum_i^N \left| f_i \right|^2$$

Equation 14

The construction of the density explicitly from a set of orbitals ensures that it is legal – it can be constructed from an asymmetric wavefunction.

If we also note that a significant component of the electron-electron interaction will be the classical Coulomb interaction – or Hartree energy (this is simply the second term of Equation 7 written in terms of the density);

$$V_H[r] = \frac{1}{2}\int \frac{r(\mathbf{r}_1)r(\mathbf{r}_2)}{\left|\mathbf{r}_1 - \mathbf{r}_2\right|} d\mathbf{r}_1 d\mathbf{r}_2$$

the energy functional can be rearranged as;

$$E[r] = T_s[r] + V_{ext}[r] + V_H[r] + E_{xc}[r]$$

Equation 15

Where we have introduced the *exchange-correlation functional;*

$$E_{xc}[r] = (T[r] - T_s[r]) + (V_{ee}[r] - V_H[r])$$

E_{xc} is simply the sum of the error made in using a non-interacting kinetic energy and the error made in treating the electron-electron interaction classically. Writing the functional (Equation 15) explicitly in terms of the density built from non-interacting orbitals (Equation 14) and applying the variational theorem (Equation 13) we find that the orbitals, which minimise the energy, satisfy the following set of equations;

$$\left[-\frac{1}{2}\nabla^2 + v_{ext}(\mathbf{r}) + \int \frac{r(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' + v_{xc}(\mathbf{r})\right] f_i(\mathbf{r}) = e_i f_i(\mathbf{r})$$

Equation 16

In which we have introduced a local multiplicative potential which is the functional derivative of the exchange correlation energy with respect to the density,

$$v_{xc}(\mathbf{r}) = \frac{dE_{xc}[r]}{dr}$$

Equation 17

This set of non-linear equations (the Kohn-Sham equations) describes the behaviour of non-interacting "electrons" in an effective local potential. For the exact functional, and thus exact local potential, the "orbitals" yield the exact ground state density via Equation 14 and exact ground state energy via Equation 15.

These Kohn-Sham equations have the same structure as the Hartree-Fock equations (Equation 8) with the non-local exchange potential replaced by the local *exchange-correlation potential* v_{xc}. We note at this point that the nomenclature in general use and reproduced here is very misleading. As stated above E_{xc} contains an element of the kinetic energy and is not the sum of the exchange and correlation energies as they are understood in Hartree-Fock and correlated wavefunction theories.

The Kohn-Sham approach achieves an exact correspondence of the density and ground state energy of a system consisting of non-interacting Fermions and the "real" many body system described by the Schrödinger equation. A cartoon representing this relationship is displayed in Figure 1.

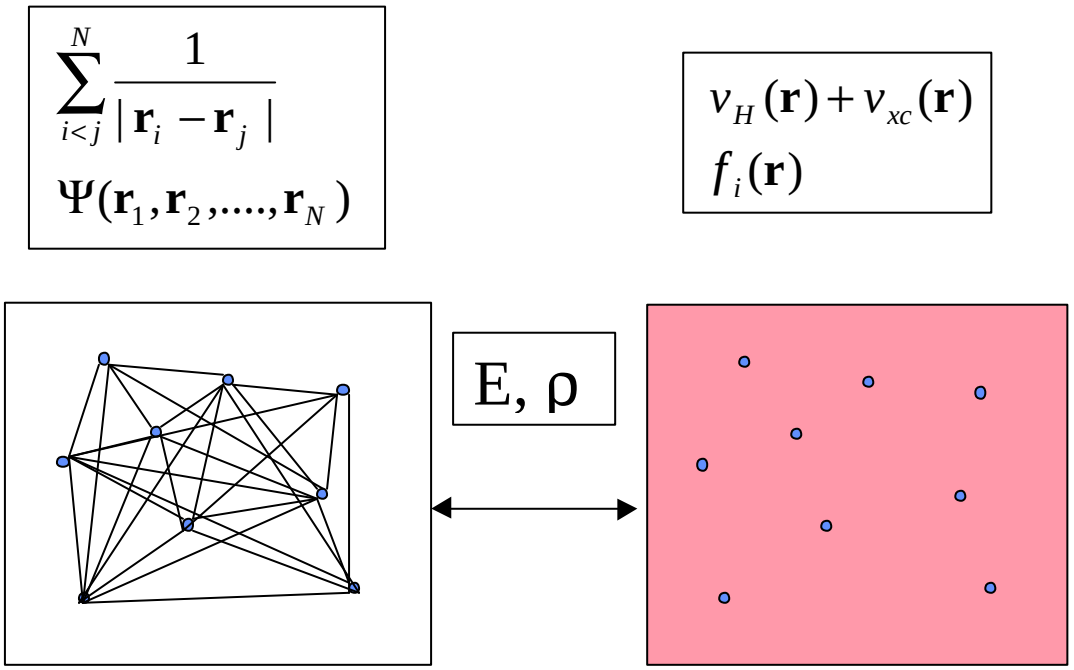


Figure 1 A cartoon representing the relationship between the "real" many body system (left hand side) and the non-interacting system of Kohn Sham density functional theory (right hand side).

The correspondence of the charge density and energy of the many-body and the non-interacting system is only exact if the exact functional is known. In this sense Kohn-Sham density functional theory is an *empirical* methodology – we do not know (and have no way of systematically approaching) the exact functional. However, the functional is universal – it does not depend on the materials being studied. For any particular system we could, in principle, solve the Schrödinger equation exactly and determine the energy functional and its associated potential. This, of course, involves a greater effort than a direct solution for the energy. Nevertheless, the ability to determine exact properties of the universal functional in a number of systems allows excellent approximations to the functional to be developed and used in *unbiased* and thus *predictive* studies of a wide range of materials – a property usually associated with an *ab initio* theory. For this reason the approximations to density functional theory discussed below are often referred to as *ab initio* or *first principles* methods.

The computational cost of solving the Kohn Sham equations (Equation 16) scales formally as N^3 (due to the need to maintain the orthogonality of N orbitals) but in current practice is dropping towards N^1 through the exploitation of the locality of the orbitals.

For calculations in which the energy surface is the quantity of primary interest DFT offers a practical and potential highly accurate alternative to the wavefunction methods discussed above. In practice, the utility of the theory rests on the approximation used for $E_{xc}[\rho]$.

5. The Local Density Approximation for $E_{xc}[r]$

The generation of approximations for E_{xc} has lead to a large and still rapidly expanding field of research. There are now many different flavours of functional available which are more or less appropriate for any particular study. Ultimately such judgments must be made in terms of results (i.e.: the direct comparison with more accurate theory or experimental data, which will be discussed below) but knowledge of the derivation and structure of functionals is very valuable when selecting which to use in any particular study.

The early thinking that lead to practical implementations of density functional theory was dominated by one particular system for which near exact results could be obtained – the homogeneous electron gas. In this system the electrons are subject to a constant external potential and thus the charge density is constant. The system is thus specified by a single number - the value of the constant electron density $\rho=N/V$.

Thomas and Fermi studied the homogeneous electron gas in the early 1920's [12]. The orbitals of the system are, by symmetry, plane waves. If the electron-electron interaction is approximated by the classical Hartree potential (that is exchange and correlation effects are neglected) then the total energy functional can be readily computed [12]. Under these conditions the dependence of the kinetic and exchange energy (Equation 7) on the density of the electron gas can be extracted (Dirac [13,1,14]) and expressed in terms of a *local* functions of the density. This suggests that in the inhomogeneous system we might approximate the *functional* as an integral over a local *function* of the charge density. Using the kinetic and exchange energy densities of the non-interacting homogeneous electron gas this leads to;

$$T[r] = 2.87\int r^{5/3}(\mathbf{r})d\mathbf{r}$$

and,

$$E_x[r] = 0.74\int r^{4/3}(\mathbf{r})d\mathbf{r}$$

Equation 18

These results are highly suggestive of a representation for E_{xc} in an inhomogeneous system. The local exchange correlation energy per electron might be

approximated as a simple function of the local charge density (say, $\varepsilon_{xc}(\rho)$). That is, an approximation of the form;

$$E_{xc}[r] \approx \int r(\mathbf{r}) e_{xc}(r(\mathbf{r}))d\mathbf{r}$$

Equation 19

An obvious choice is then to take $\varepsilon_{xc}(\rho)$ to be the exchange and correlation energy density of the uniform electron gas of density ρ - this is the *local density approximation* (LDA). Within the LDA $\varepsilon_{xc}(\rho)$ is a *function* of only the local value of the density. It can be separated into exchange and correlation contributions;

$$e_{xc}(r) = e_x(r) + e_c(r)$$

Equation 20

The Dirac form can be used for ε_x (Equation 18);

$$e_x(r) = -Cr^{1/3}$$

Equation 21

Where for generality a free constant, C, has been introduced rather than that determined for the homogeneous electron gas. This functional form is much more widely applicable than is implied from its derivation and can be established from scaling arguments [1]. The functional form for the correlation energy density, $\varepsilon_{c,}$ is unknown and has been simulated for the homogeneous electron gas in numerical quantum Monte Carlo calculations which yield essentially exact results [15]. The resultant exchange correlation energy has been fitted by a number of analytic forms [16,17,18] all of which yield similar results in practice and are collectively referred to as LDA functionals.

A cartoon of the principles underlying the LDA is displayed in Figure 2. The energy density for the material represented by the inhomogeneous density of the left hand panel, is assigned a value from the known density variation of the exchange-correlation energy density of the homogeneous electron gas (right hand panel). This value is assigned with no regard to location or variations of the inhomogeneous density.

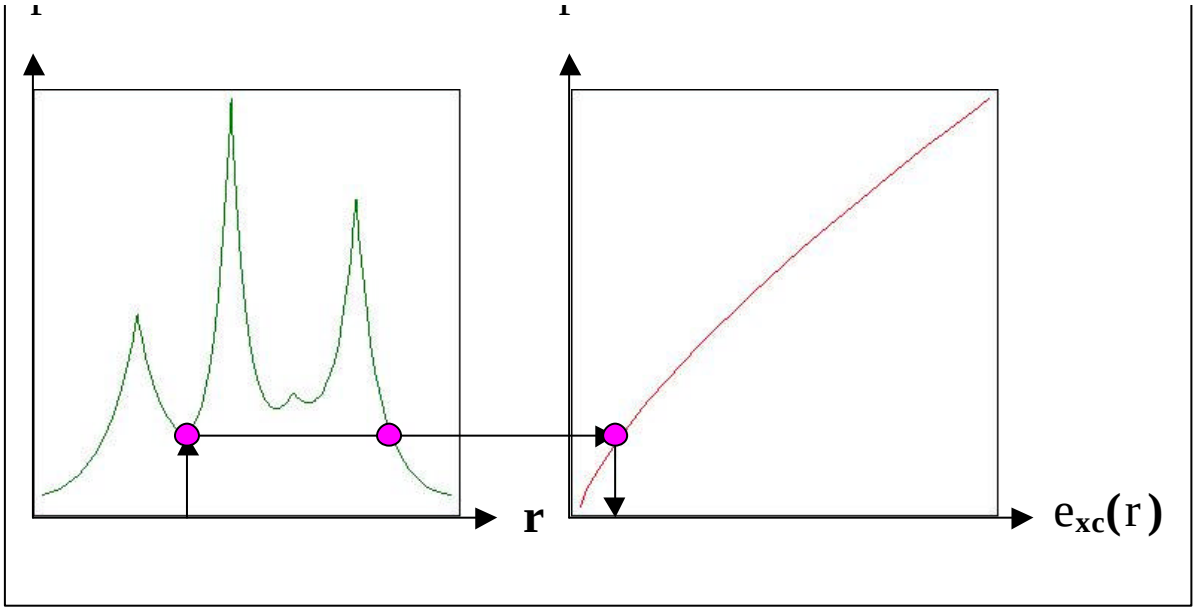


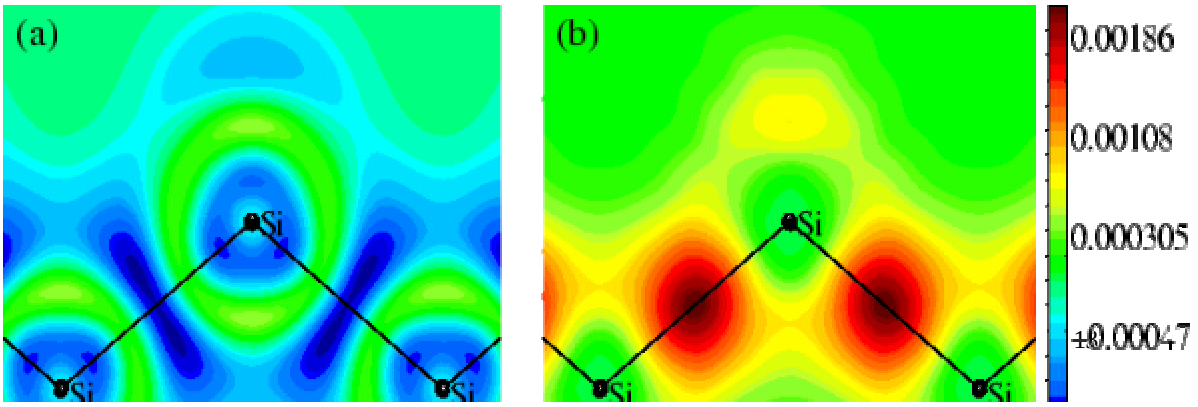
Figure 2 A cartoon depicting the use of the local density approximation to determine the exchange correlation energy of an inhomogeneous system. The exchange energy density of the homogeneous electron gas is depicted on the right and the density of the inhomogeneous system on the left [19].

The LDA has proven to be a remarkably fruitful approximation. Properties such as structure, vibrational frequencies, elastic moduli and phase stability (of similar structures) are described reliably for many systems (explicit cases will be discussed below). However, in computing energy differences between rather different structures the LDA can have significant errors. For instance, the binding energy of many systems is overestimated (typically by 20-30 %) and energy barriers in diffusion or chemical reactions may be too small or absent. Nevertheless, the remarkable fact is that the LDA works as well as it does given the reduction of the energy functional to a simple local function of the density.

Figure 3 The difference between the exchange (left) and correlation (right) energy densities computed using variation quantum Monte Carlo and the local density approximation in bulk Silicon[20].

The magnitude of the errors in the LDA energy densities has recently been estimated by computing the energy density of bulk silicon with variational quantum Monte Carlo (V-QMC) calculations. In Figure 3 the difference between LDA densities and the, V-QMC result is plotted. There are very significant errors in the exchange and correlation energies but, as the exchange energy is generally underestimated and the correlation energy overestimated, these errors tend to cancel. The success of the LDA appears to be in part due to this cancellation of errors.

For further insight into the reasons for the success of the LDA we return to our discussion of the density matrices. As demonstrated above (Equation 12) the first and



second order density matrices are sufficient to determine the exact total energy. Insight into the behaviour of functionals can be obtained by examining how well they approximate P_2. A commonly used device is to convert P_2 - the probability of finding an electron at $\mathbf{r}_1$ and an electron at $\mathbf{r}_2$ – into the conditional probability of finding an electron at $\mathbf{r}_2$ given that there is an electron at $\mathbf{r}_1$; this quantity is the exchange correlation hole;

$$P_{xc}(\mathbf{r}_1, \mathbf{r}_2) = \frac{P_2(\mathbf{r}_1, \mathbf{r}_2)}{r(\mathbf{r}_1)} - r(\mathbf{r}_2)$$

Equation 22

We may think of P_{xc} as the hole an electron placed at $\mathbf{r}_1$ digs for itself in the surrounding density. There are a number of exact properties of the hole which one would hope to reproduce in approximations. For instance, for any point $\mathbf{r}_1$ the reduction in the surrounding density should be one-electron, that is;

$$\int P_{xc}(\mathbf{r}_1, \mathbf{r}_2) d\mathbf{r}_2 = -1$$

Equation 23

This result follows immediately if P_2 of Equation 22 is inserted in Equation 11.

The exact variation of the exchange hole for a neon atom is plotted as a function of $\mathbf{r}_2$ for $\mathbf{r}_1$=0.09a_0 and 0.4 a_0 Figure 4 and compared to that computed within the LDA (from ref. [21]). It is clear that the LDA is a very poor approximation to P_2.

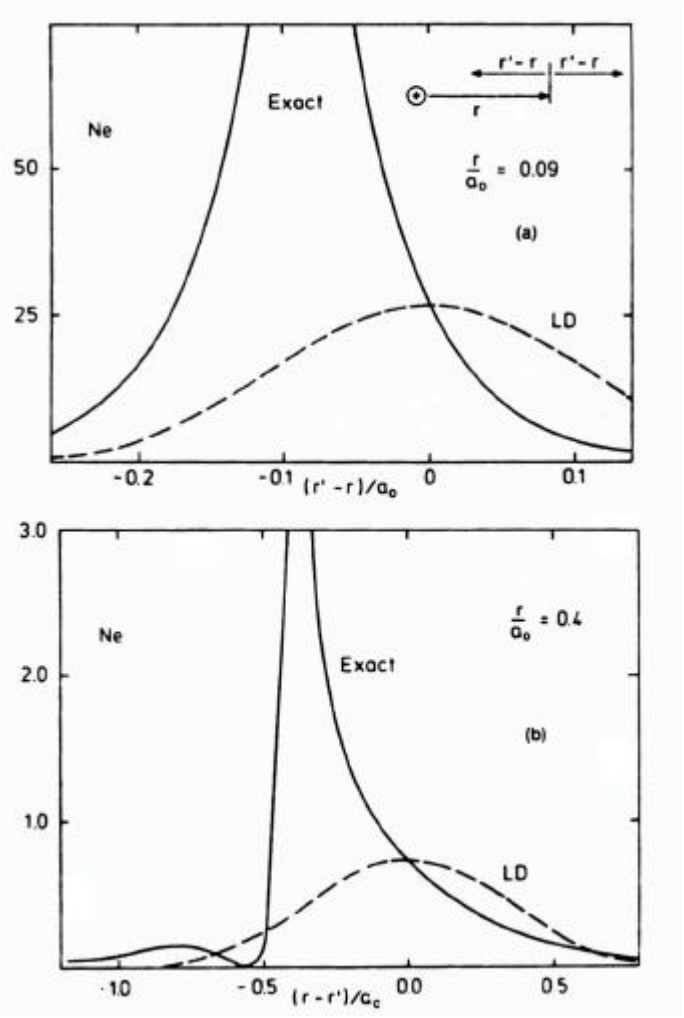


Figure 4 The exchange hole, $P_x(\mathbf{r},\mathbf{r}')$, for a neon atom comparing the exact result (continuous line) to that of the LDA (dashed line) [21]. The top panel is for r=0.09a_0 and the lower for r=0.4a_0

We are faced with the question – how can the LDA produce such reasonable energetics if the pair correlation function is so poorly described? The answer is based on the structure of the Coulomb operator. We remember from Equation 12 that the electron-electron interaction can be written in terms of P_2 as;

$$V_{ee} = \frac{1}{2}\int \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} P_2(\mathbf{r}_1, \mathbf{r}_2)\, d\mathbf{r}_1 d\mathbf{r}_2$$

Equation 24

From this it seems readily apparent that a poor approximation to P_2 leads directly to a poor estimate of the electron-electron interaction. However, the Coulomb operator depends only on the magnitude of the separation of $\mathbf{r}_1$ and $\mathbf{r}_2$ Substitution of $\mathbf{u} = \mathbf{r}_1 - \mathbf{r}_2$ yields;

$$V_{ee} = \frac{1}{2}\int\left[\int\frac{1}{u}P_2(\mathbf{r}_1,\mathbf{r}_1+\mathbf{u})d\mathbf{r}_1\right]d\mathbf{u}$$
$$= \frac{1}{2}\int_0^\infty 4pu^2.\left[\frac{\int P_2(\mathbf{r}_1,\mathbf{r}_1+\mathbf{u})d\mathbf{r}_1}{u}\frac{d\Omega_u}{4p}\right]du$$

Equation 25

Thus the electron-electron interaction depends only on the *spherical average* of the pair density - $P(u)$;

$$.P(u) = \int P_2(\mathbf{r}_1,\mathbf{r}_1+\mathbf{u})d\mathbf{r}_1\frac{d\Omega_u}{4p}$$

Equation 26

The exact P(u) for the neon atom is compared to that resulting from the LDA in Figure 5 for the same positions of **r** used in Figure 4. The LDA makes a reasonable approximation to the spherically averaged hole and preserves the normalisation of the hole to –1 (Equation 23). This observation explains in part the success of the LDA.

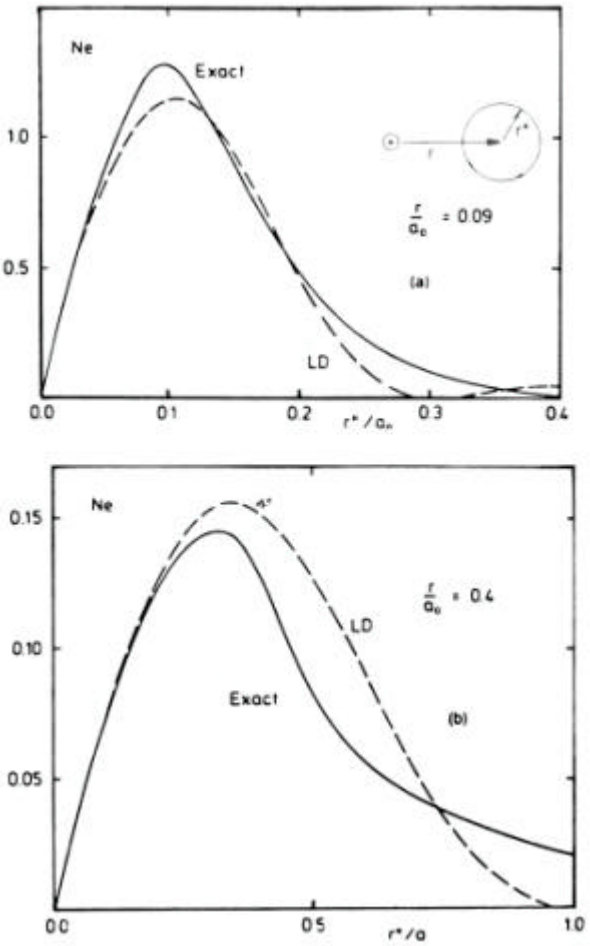


Figure 5 The spherical average of the exchange hole in a neon atom. comparing the exact result (continuous line) with that computed in the LDA (dashed line) [21].

We can conclude that the remarkable performance of the LDA is a consequence of its reasonable description of the spherically averaged exchange correlation hole coupled with the tendency for errors in the exchange energy density to be cancelled by errors in the correlation energy density. An understanding of these features is an important pre-requisite to developing functionals that seek to improve on the LDA

6. Beyond the Local Density Approximation

At first sight a very natural extension of the LDA would be to recognise that in many systems the exchange contribution to the energy is dominant over the correlation energy and to compute the non-local exchange potential exactly as in Hartree Fock theory (Equation 7) whilst approximating the correlation potential within the LDA. This would yield a functional of the form:

$$E_{xc} \approx E_{Fock} + E_c^{LDA}$$

The greater complexity [22] associated with the calculation of the non-local exchange potential would be offset by potentially significantly greater accuracy. However, the performance of the LDA is, in part, based on rather delicate cancellations between the exchange and correlation interactions and, in general, the use of the exact exchange interaction produces rather poor results.

In the homogeneous electron gas the non-local exchange potential has effectively infinite range and its contribution to the electron-electron interaction diverges at the Fermi surface [23]. In metals we conclude that the non-local exchange potential does not yield the correct physics – indeed this behaviour was one of the main motivating factors in the early work of Thomas and Fermi [12], which formed the basis of density functional theory.

Further evidence for the poor behaviour of the non-local exchange interaction in systems where the band gap (HOMO-LUMO gap) is small is obtained from studies of the H_2 molecule at extended bond lengths. The exchange correlation hole for this system is displayed in Figure 6.

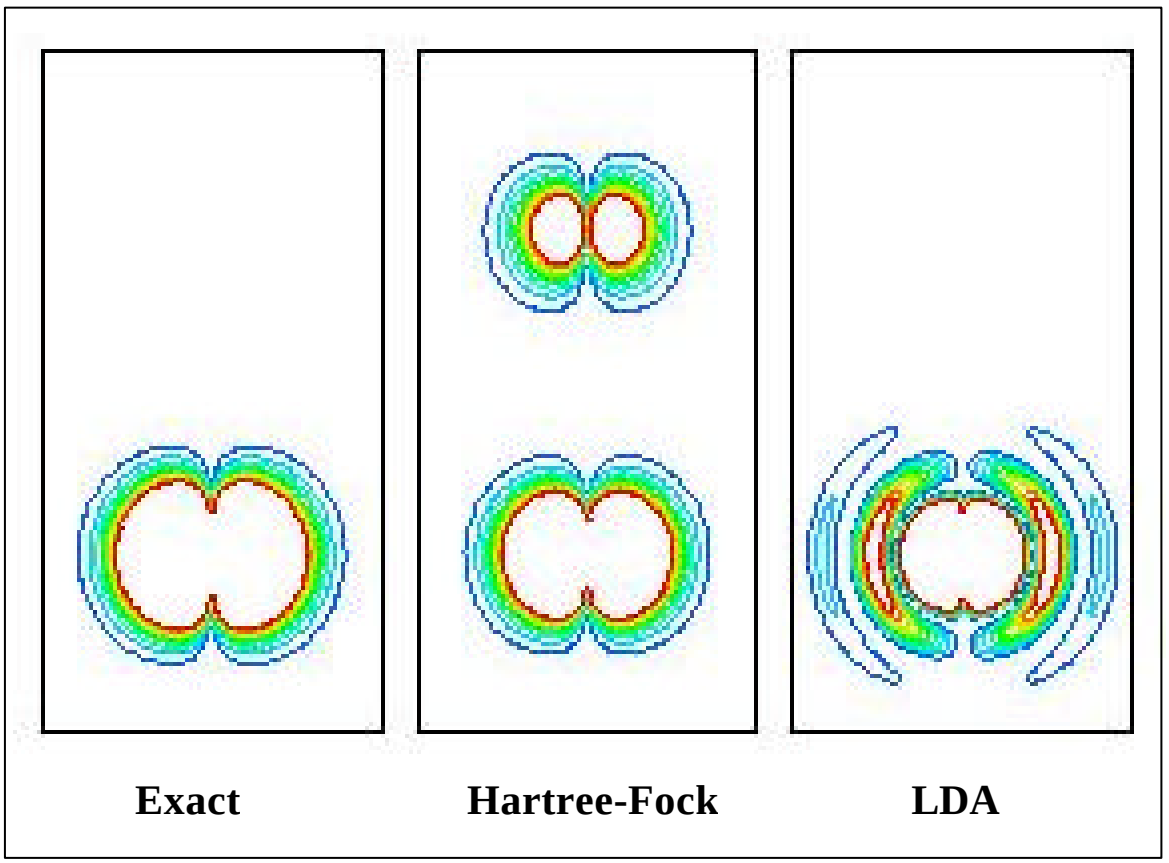


Figure 6 The exchange correlation hole in H_2 at an extended bond lengths computed exactly, with the non-local exchange potential of Hartree-Fock theory and in the LDA. The hole density is plotted as a function of r' with r centred on the H_2 bond [19].

The Hartree-Fock potential produces a very reasonable local and semi-local description – which is superior to that of the LDA – but introduces a pathological non-local feature.

We can conclude that in order to improve on the LDA approximation semi-local theories which incorporate some of the features of the exact exchange interaction are required and that theories which preserve the analytic properties of the exchange correlation hole are likely to be successful.

7. The Generalised Gradient Approximation

The local density approximation can be considered to be the zeroth order approximation to the semi-classical expansion of the density matrix in terms of the

density and its derivatives [24]. A natural progression beyond the LDA is thus to the gradient expansion approximation (GEA) in which first order gradient terms in the expansion are included. This results in an approximation for the exchange hole [24] which has a number of unphysical properties; it does not normalise to –1, it is not negative definite and it contains oscillations at large u [25].

In the generalised gradient approximation (GGA) a functional form is adopted which ensures the normalisation condition and that the exchange hole is negative definite [26,27]. This leads to an energy functional that depends on both the density and its gradient but retains the analytic properties of the exchange correlation hole inherent in the LDA.

The typical form for a GGA functional is;

$$E_{xc} \approx \int r(r) e_{xc}(r, \nabla r) d\mathbf{r}$$

Equation 27

As we will see below the GGA improves significantly on the LDA's description of the binding energy of molecules – it was this feature which lead to the very wide spread acceptance of DFT in the chemistry community during the early 1990's. A number of functionals within the GGA family [28,26,27,29,30,31] have been developed. The performance of these functionals will be discussed below.

8. Meta-GGA functionals

Recently functionals that depend explicitly on the semi-local information in the Laplacian of the spin density or of the local kinetic energy density have been developed [32,33,34]. Such functionals are generally referred to as meta-GGA functionals.

The form of the functional is typically;

$$E_{xc} \approx \int r(r) e_{xc}(r, |\nabla r|, \nabla^2 r, t) d\mathbf{r}$$

Equation 28

Where the kinetic energy density τ is;

$$t = \frac{1}{2} \sum_i |\nabla j_i|^2$$

Equation 29

9. Hybrid Exchange Functionals

There is an *exact* connection between the non-interacting density functional system and the fully interacting many body system via the integration of the work done in gradually turning on the electron-electron interactions. This adiabatic connection approach [35] allows the exact functional to be formally written as:

$$E_{xc}[r] = \frac{1}{2}\int d\vec{\mathbf{r}} d\vec{\mathbf{r}}' \int_{l=0}^{1} dl \, \frac{le^2}{|\vec{\mathbf{r}} - \vec{\mathbf{r}}'|}[< r(\vec{\mathbf{r}}) r(\vec{\mathbf{r}}') >_{r,l} - r(\vec{\mathbf{r}}) d(\vec{\mathbf{r}} - \vec{\mathbf{r}}')]$$

Equation 30

where the expectation value $< ... >_{r,l}$ is the density-density correlation function and is computed at density $r(\mathbf{r})$ for a system described by the effective potential;

$$V_{eff} = V_{en} + \frac{1}{2}\sum_{i \neq j} \frac{le^2}{|\vec{\mathbf{r}}_\mathbf{i} - \vec{\mathbf{r}}_\mathbf{j}|}$$

Equation 31

Thus the exact energy could be computed *if* one knew the variation of the density-density correlation function with the coupling constant, λ. The LDA is recovered by replacing the pair correlation function with that for the homogeneous electron gas.

The adiabatic integration approach suggests a different approximation for the exchange-correlation functional. At λ=0 the non-interacting system corresponds identically to the Hartree-Fock ansatz, while the LDA and GGA functionals are constructed to be excellent approximations for the fully interacting homogeneous electron gas – that is, a system with λ=1. It is therefore not unreasonable to approximate the integral over the coupling constant as a weighted sum of the end points – that is, we might set:

$$E_{xc} \approx aE_{Fock} + bE_{xc}^{GGA}$$

with the coefficients are to be determined by reference to a system for which the exact result is known. Becke adopted this approach [36] in the definition of a new functional with coefficients determined by a fit to the observed atomisation energies, ionisation potentials, proton affinities and total atomic energies for a number of small molecules [36]. The resultant (three parameter) energy functional is,

$$E_{xc} = E_{xc}^{LDA} + 0.2(E_X^{Fock} - E_X^{LDA}) + 0.72\Delta E_X^{B88} + 0.81\Delta E_c^{PW91}$$

Here ΔE_X^{B88} and ΔE_c^{PW91} are widely used GGA corrections [37,38] to the LDA exchange and correlation energies respectively.

Hybrid functionals of this type are now very widely used in chemical applications with the B3LYP functional (in which the parameterisation is as given above but with a different GGA treatment of correlation [39]) being the most notable. Computed binding energies, geometries and frequencies are systematically more reliable than the best GGA functionals.

10. The Performance of Various Functionals

The functionals currently used in density functional simulations form a natural hierarchy. Although it cannot be claimed that there is a systematic approach to the exact functional it is clear that improvements are being made in the underlying functional form and that the description of ground state properties is improving. The most notable recent advances being those in which the non-local nature of the exchange potential is introduced in one form or another. The current hierarchy is summarized in Table 1.

Dependencies	**Family**
Exact exchange, $\|\nabla\rho\|$, ρ	Hybrid
$\nabla^2\rho$, τ	Meta-GGA
$\|\nabla\rho\|$	GGA
ρ	LDA

Table 1 The current hierarchy of exchange correlation functionals.

In producing functionals there are two broad schools of thought, which may be summarized as follows;

- Adopt an appropriate functional form and introduce parameters to be determined by reference to experimental data or data from explicitly correlated calculations. This is largely an empirical approach.
- Use the exact properties of the functional to determine both its structure and the parameters in its functional form.

There are, of course, a number of functionals that rather cross the boundary between the two schools but, as we will see below, the distinction is often valuable in assessing the likely accuracy of a particular functional in a new application. Unfortunately in order to discuss a variety of functionals they must be labeled with a mnemonic – a selection of which are listed in Table 2 along with an indication of the family they come from and the number of empirically determined parameters which enter the functional. The parameters are determined in practice from fitting to the properties of large training sets of molecular properties.

Mnemonic	Family	Parameters
LDA	Local	-
BLYP	GGA	Light
PBE	GGA	-
HCTH	GGA	18
VS98	Meta-GGA	21
PKZB	Meta-GGA	1
Hybrid	Hybrid-exchange	3

Table 2 Mnemonics used to define density functionals of various families and the number of external parameters fitted when defining the functional.

Apart from the LDA the functionals listed are the GGA functionals BLYP [29, 39], PBE [31], HCTH [40] the meta-GGA functionals VS98 [41], PKZB [34] and the Hybrid functional of Becke [36]. The degree of empirical parameterisation of these functionals varies considerably with the BLYP, PBE and PKZB functionals aiming to be virtually *ab-initio* and parameter free while the HCTH and VS98 functionals are very heavily parameterised with reference to large molecular training sets.

The performance on these functionals in calculations of a number of molecular and material properties have recently been tabulated by Kurth *et. al.* [42] and Adamo *et. al.* [43]. Here we summarise some of the key data from these studies.

Atomisation Energies		
	M.R.E %	**M.A.E. (max) kcal/mol**
LDA	22 %	-
BLYP	5 %	-
PBE	7 %	17 (51)
HCTH	3 %	-
VS98	2 %	3 (12)

PKZB	3 %	5 (38)
Hybrid	-	3 (20)

Table 3 Atomisation Energies - the mean relative error (M.R.E) for a collection of 20 molecules [42] **and the mean absolute error (M.A.E.) in kcal/mol for a collection of 148 molecules** [43] **with the maximum absolute error given in brackets.**

In Table 3 errors in computed molecular atomisation energies are tabulated. The tendency of the LDA to over bind by 20-30% discussed above is seen clearly in this data. The GGA functionals yield very significant improvements with relative errors in the range 3-7% and an average absolute error of 17 kcal/mol. The highly parameterised HCTH functional performs somewhat better than the BLYP and PBE functionals. The meta-GGA functionals yield relative errors of 2-3% and average absolute errors of 3-5 kcal/mol. Again the highly parameterised functional – VS98 - performs slightly better than PKZB. The meta-GGA functionals achieve a distinct improvement over the GGA functionals. The relatively lightly parameterised hybrid functional performs as well as the best meta-GGA. For many chemical reactions an accuracy of 1-2 kcal/mol is sufficient for predictive modelling of the thermochemistry. It is clear that average errors yielded by the meta-GGA and hybrid functionals are approaching this limit. Performance is not systematically at this level; maximum errors for 'difficult' systems are still in the range 12-38 kcal/mol.

Structures		
	M.R.E. % [42]	**M.A.E. (max) Angstrom [43]**
LDA	5%	-
BLYP	8%	-
PBE	4%	0.011 (0.064)
HCTH	6%	-
VS98	8%	0.008 (0.08)
PKZB	3%	0.019 (0.111)
Hybrid	-	0.007 (0.062)

Table 4 Mean relative error (M.R.E.) in the unit cell volumes of 12 crystals [42] **and mean absolute errors (M.A.E.) in the bond lengths of 23 molecules** [43] **with the maximum absolute error given in brackets.**

Results for unit cell volumes of 12 different crystal structures and for the bond lengths in 23 molecules are summarised in Table 4. The LDA is a remarkably good approximation in bulk crystals and its error for volumes (5%) compares favourably to that produced by GGA functionals (4-8%) and meta-GGA functionals (3-8%). In molecular systems the highly parameterised meta-GGA and the hybrid functional are somewhat more accurate than the PBE GGA functional. It is interesting to note that although the VS98 functional is somewhat more accurate than the PKZB in molecules this situation is reversed for calculations on crystals. This is part of a general trend in that lightly parameterised functionals do not contain a bias towards molecular systems

and tend to be transferable to materials applications while highly parameterised functionals are not.

Bulk Modulus and Vibrational Frequencies		
	Bulk Modulus M.R.E. %	**Vibrational Frequencies M.A.E. (max) (cm^{-1})**
LDA	19%	-
BLYP	22%	-
PBE	10%	65 (-194)
HCTH	20%	-
VS98	29%	33 (-109)
PKZB	9%	72 (+144)
Hybrid	-	40 (-209)

Table 5 Mean relative error (M.R.E) in the bulk moduli of 12 crystals [42] **and mean absolute error (M.A.E.) in the vibrational frequencies of 55 molecules** [43] **with the maximum absolute error given in brackets.**

Results for the bulk moduli of 12 different crystal structures and for the vibrational frequencies in a set of 55 molecules are summarised in Table 5. For the calculation of bulk moduli the GGA functionals offer a small but non-systematic improvement over the 19% error of the LDA. The PKZB meta-GGA functional produces a significantly smaller error of 9%. For vibrational frequencies in molecules the VS98 and hybrid exchange functionals perform somewhat better than the PBE GGA functional or the PKZB meta-GGA functional. As for computed structures there is a tendency for the highly parameterised VS98 meta-GGA functional to perform well in molecules but rather poorly in crystalline solids.

11. Conclusions

Density functional theory provides us with a relatively efficient and *unbiased* tool with which to compute the ground state energy in realistic models of bulk materials and their surfaces. The reliability of such calculations depends on the development of approximations for the exchange-correlation energy functional. Significant advances have been made in recent years in the quality of exchange correlation functionals as dependence on local density gradients, semi-local measures of the density and non-local exchange functionals have been introduced.

The local density approximation is a very simple and remarkably reliable for the structure, elastic moduli, relative phase stability of many materials but is less accurate for binding energies and details of the energy surface away from equilibrium geometries – eg: transition states. The GGA family of functionals improves binding energies to average errors of 20 kcal/mol and relative errors of 3-7% while meta-GGA

and hybrid-exchange functionals reduce these errors to 3-5 kcal/mol and 2-3%. This is close to the accuracy required for predictive simulations of thermochemical properties. The GGA, meta-GGA and hybrid functionals retain, and somewhat improve on, the LDA's excellent description of bonds lengths with typical errors in the region of 1-2 milli-Angstrom. Using these functionals elastic moduli are reproduced to within 10% and vibrational frequencies to ~40cm^{-1}.

There is a distinct tendency for functionals that are highly parameterised and fitted to the properties of molecular systems to perform somewhat better than lightly parameterised functions for molecules but to perform relatively poorly in simulations on periodic materials.

---

1 R. G. Parr, and W. Yang, Density-Functional Theory of Atoms and Molecules, OUP, Oxford, (1989)
2 N. H. March (ed.), Electron Correlation in the Solid State, ICP, London, (1999)
3 J. Callaway and N. H. March, Density Functional Methods: Theory and Applications, *Solid State Physics*, **38**, 135 (1984)
4 A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry*, (Macmillan, New York, 1982)
5 B. L. Hammond and W. A. Lester Jr. and P.J. Reynolds, *Monte Carlo Methods in Ab Initio Quantum Chemistry*, (World Scientific, Singapore, 1994).
6 J. Coleman, Calculation of the first- and second-order density matrices, in, The Force Concenpt in Chemistry, Ed. B. M. Deb, (Van Nostrand Reinhold, New York, 1981)
7 R. Erdahl and V. H. Smith Jr. Eds. *Density matrices and Density Functionals*, (Reidel, Dordrecht 1987).
8 P. Hohenburg and W. Kohn, *Phys. Rev.* **136** B864 (1964)
9 M. Levy, *Proc. Natl. Acad. Sci. USA* **76**, 6062 (1979); M. Levy, *Phys. Rev. A* **26**, 1200 (1982); M. Levy and J. P. Perdew, *The Constrained Search Formalism for Density Functional Theory*, in *Density Functional Methods in Physics*, Ed. R. M. Dreisler and J. da Providencia, (Plenum, New York 1985)
10 M. Foley and P. A. Madden, *Phys. Rev.* B **53**, 10589 (1996)
11 W. Kohn and L. J. Sham, *Phys. Rev.* **140** A1133 (1956)
12 E. Fermi *Z. Phys.* **48** 73 (1928); L. H. Thomas, *Proc. Camb. Phil. Soc.* **23** 542 (1927); these articles are reproduced in N. H. March, Self Consistent Fields in Atoms, Plenum, Oxford, (1975).
13 P. A. M. Dirac, *Proc. Camb. Phil. Soc.* **26** 376 (1930).
14 E. H. Lieb, *Rev. Mod. Phys.*, **53** 603 1981.
15 D. M. Ceperley and B. J. Alder, *Phys. Rev. Lett.*, **45** 566 1980
16 J. P. Perdew and A. Zunger, *Phys. Rev.* B **23** 5048 (1981)
17 U. von Barth and L. Hedin, *J. Phys.* C **5** 1629 (1972)
18 S. H. Vosko, L. Wilk and M. Nusair, Accurate spin-dependent electron liquid correlation energies of local spin density calculations: a critical analysis", *Can. J. Phys.* **58**, 1200 (1980)
19 This image is used with the kind permission of A. Savin (unpublished work).

20 R. Q. Hood, M. Y. Chou, A. J. Williamson, R. Rajagopal, R.J. Needs and W. M. C. Foulkes, *Phys. Rev. Lett.* **78** 3350 1997.
21 O. Gunnarsson, M. Jonson and B. I. Lundqvist, *Phys. Rev.* B **20** 3136 (1979).
22 The exact exchange potential can be computed rather readily for periodic materials when a basis set Gaussian type orbitals is used as in the CRYSTAL software: R. Dovesi, V. R. Saunders, C. Roetti, M. Causà, N. M. Harrison, R. Orlando, C. M. Zicovich-Wilson, CRYSTAL98 User's Manual, University of Torino, Torino, 1998; (*http://www.cse.dl.ac.uk/Activity/CRYSTAL*).
23 See, for instance, the discussion in Chapter 17 of, *Solid State Physics*, N. W. Ashcroft and N. D. Mermin, second edition, (HoltSaunders, Philadelphia, 1976)
24 R. M. Dreizler and E. K. U. Gross, *Density Functional Theory* (Springer Verlag, Berlin, 1990)
25 Y. Yang, J. P. Perdew, J. A. Cevary, L. D. Macdonald and S. H. Vosko, *Phys. Rev.* A **41**, 78 (1990)
26 J. P. Perdew and Y. Wang, *Phys. Rev.* B **33**, 8800, (1986); Ibid. E **34**, 7406, (1986)
27 J. P. Perdew, in *Electronic Structure of Solids 91*, Ed. P. Ziesche and H. Eschrig (Akademie Verlay, Berlin, 1991)
28 D. C. Langreth, M. J. Mehl, *Phys. Rev.* B, **28**, 1809 (1983)
29 A. D. Becke, *Phys. Rev.* A **38** 3098 1988
30 C. Lee, W. Yang, R. G. Parr, *Phys. Rev.* B **37** 785 (1988)
31 J. P. Perdew, K. Burke, M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865, (1996); Ibid. E **78,** 1396 (1997)
32 V. Tschinke, T. Zieglar, *Can. J. Chem.* **67**, 460, (1989)
33 R. Neumann, N. C. Handy, Chem. Phys. Lett.,**266**, 16 (1997)
34 J. P. Perdew, S. Kurth, A. Zupan and P. Blaha, *Phys. Rev. Lett.* **82**, 2544, (1999); Ibid E **82**, 5179 (1999)
35 D. Pines and P. Nozières, *The Theory of Quantum Liquids* (Benjamin, New York, 1966); J. Harris and R. O. Jones, *J. Phys.* F **4,** 1170 (1974), O. Gunnarsson and B. I. Lundqvist, *Phys. Rev.* B **13**, 4274 (1976).
36 A. D. Becke, *J. Chem. Phys*., **98**, 1372 (1993); A.D. Becke, Ibid. **98,** 5648 (1993)
37 A. D. Becke, *J. Chem. Phys.*, **88**, 1053 (1988)
38 J. P. Perdew and Y. Wang, *Phys. Rev.* B **45,** 13244 (1992)
39 C. Lee, W. Yang and R. G. Parr, *Phys. Rev.* B **37**, 785 (1988)
40 F. A. Hamprecht, A. J. Cohen, D. J. Tozer and N. C. Handy, *J. Chem. Phys.* **109**, 6264 (1998)
41 T. Van Voorhis and G. E. Scuseria, *J. Chem. Phys.* **109**, 400, (1998)
42 S. Kurth, J. P. Perdew and P. Blaha, *Int. J. of Quant. Chem.* **75** 889 (1999)
43 C. Adamo, M. Ernzerhof and G. E. Scuseria, *J. Chem. Phys.*, **112**, 2643 (2000)