Density functional theory

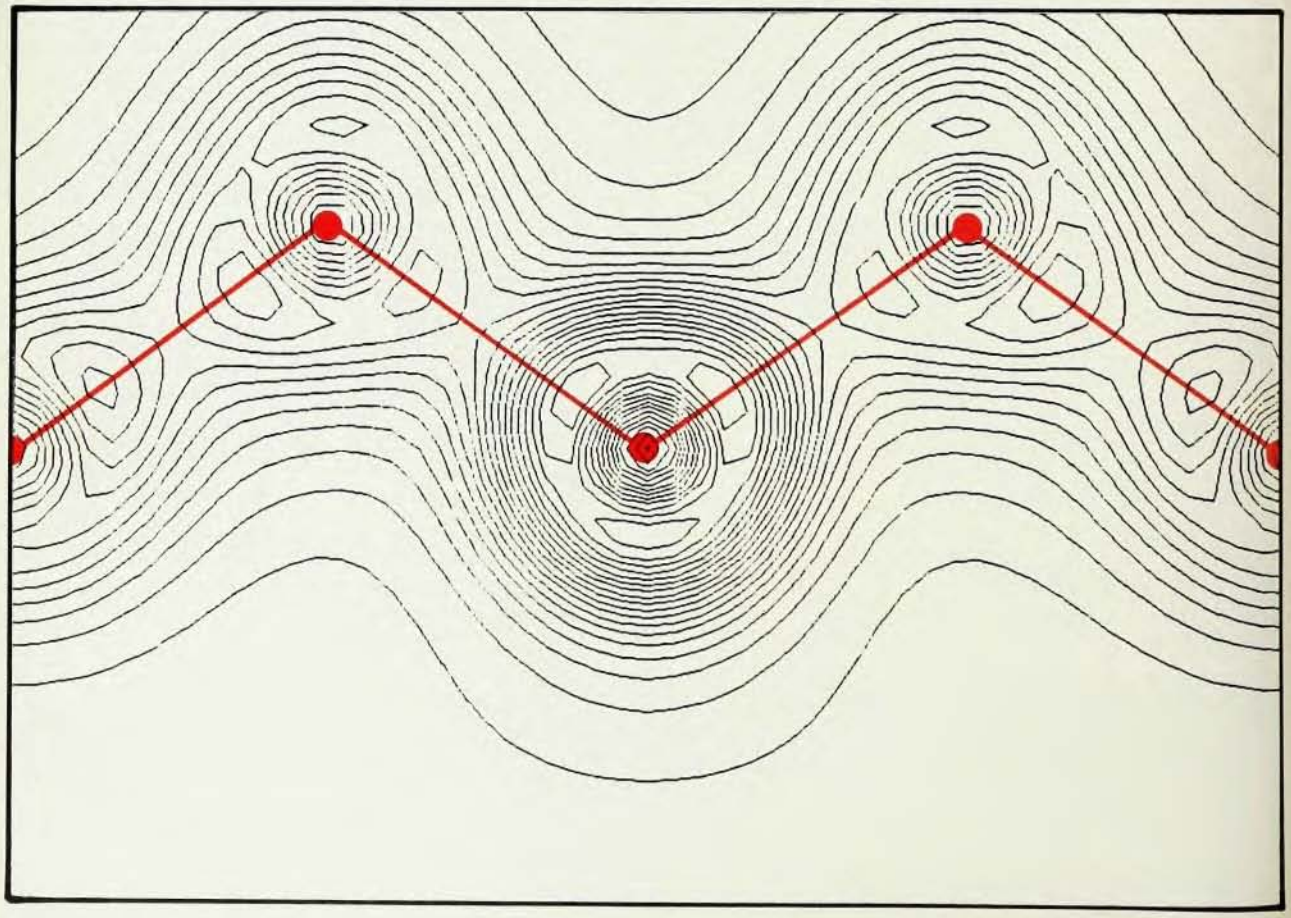
A scheme for incorporating electron density distributions into calculations on many-particle systems yields good predictions of quantities such as binding energy in molecules and phonon spectra in solids.

Michael Schlüter and Lu Jeu Sham

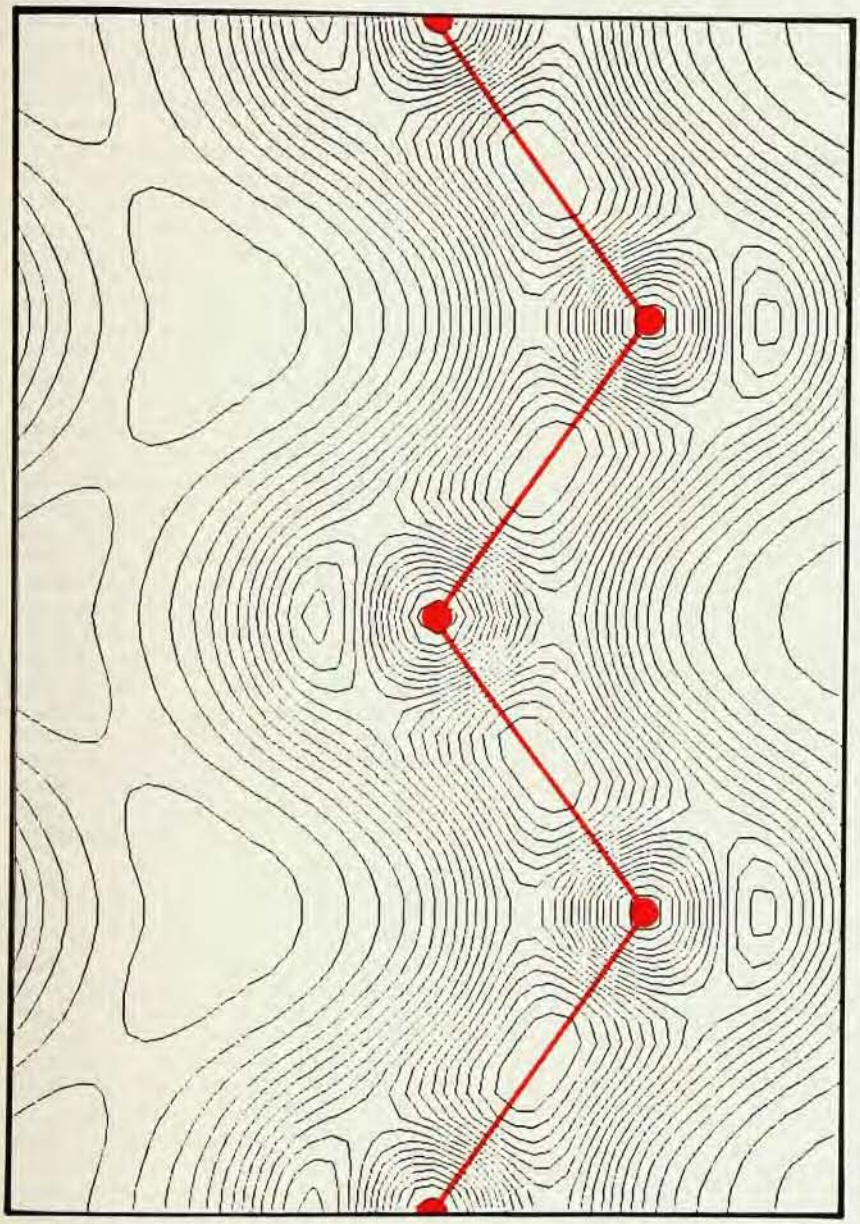
What are the energies and wavefunctions of electrons under the influence of nuclei as well as other electrons? If we could solve this general theoretical problem, we would gain a fundamental understanding of a healthy chunk of atomic, molecular and solid-state physics.

If the nuclei are arranged on a periodic lattice, for example, knowledge of the electronic states allows us to calculate thermal, optical and magnetic properties of the solid, equations of state, electron density distributions (such as those shown in figure 1) and cohesive energies, all of which we can compare with observation. When the lattice is slightly distorted to simulate the atomic arrangement in the presence of a phonon, the solution to our problem allows us to predict the lattice vibrational spectrum. And when we put a metal's electron and phonon properties together, we can calculate its superconducting transition temperature. Moreover, even though we exclude excited-state properties such as heat transport as beyond the scope of our problem, its

Michael Schlüter is a member of the technical staff at Bell Laboratories, Murray Hill, New Jersey. Lu Jeu Sham is professor of physics at the University of California, San Diego.



0031-9228 / 82 / 020036-08 / $01.00



Calculated distribution of charge as it arises from chemically active valence electrons. The contours of equal charge density show a nitrogen impurity atom in an otherwise perfect diamond crystal (left) and a perfect silicon crystal (above). The nitrogen atom, being more electronegative than its diamond host, attracts extra electrons. The bonding charge between pairs of carbon atoms and between carbon and nitrogen atoms shows a characteristic double-peak structure, indicated on the left by pairs of closed loops between atoms. This is not found in other tetrahedral semiconductors such as silicon, whose covalent bonds show a single concentration of charge between atoms. Large interstitial regions are virtually empty, creating a periodic network of electronic charge that is characteristic of insulators and semiconductors. In metals the charge is more evenly spread out. The calculations were done with a local-density functional method using pseudopotentials. A Green's function scattering technique was used to describe the nitrogen defect. (From reference 9.) Figure 1

solution is, nonetheless, basic to the calculation of these properties as well.

In this article we will examine a particular approach to the problem posed above, an approach that has developed over the last seventeen years into a useful tool for the calculation of electronic properties and for the general exploration of solid-state physics: the density-functional method.

Until the mid-sixties, calculations of the electronic states in solids tended to emphasize either the electron–nucleus interaction or the electron–electron interaction. Band structure calculations—solutions of the one-electron Schrödinger equation—generally emphasize the electron–nucleus interaction. Progress in obtaining accurate band structure calculations has been such that, today, different methods of solution yield answers in quantitative agreement with each other as long as we start with the same effective one-electron potential.

As a result, we can now aim questions at the adequacy of the one-electron potential in accounting for the electron–electron interaction as well as for the lattice interaction. The first-principles Hartree potential, where the electron–electron interaction is approximated by the Coulomb potential arising from the charge distribution of the electrons treated as fixed, is generally regarded as inadequate. The Hartree–Fock approximation has the Pauli principle built in explicitly, which results in what we call the exchange interaction among electrons of equal spin. Electrons of opposite spin remain unaffected. This approximation has pathological features, such as a zero density of states at the Fermi level in metals. Moreover, it is so complex in solids that few successful calculations have been done.

On the other hand, the theory of electron correlation has made great strides in the "jellium" approximation, that is, where the solid is modeled by a system of interacting electrons in a uniform positively-charged background. This approximation neglects the effect of the lattice potential. The twin approach of first-principles calculations, utilizing the methods of quantum field theory, and the phenomenological theory of Lev Landau, first constructed for the Fermi liquid helium-3, led to reasonable values for the cohesive energy, enhancement of the effective mass and specific heat (or density of states), and to knowledge about a veritable zoo of collective modes in interacting electron systems. Clearly, jellium is at best a zeroth order approximation for metals.

The next order approximation is made possible by James C. Phillips and Leonard Kleinman's theory of the pseudopotential.[1] The pseudopotential includes both the lattice potential and a repulsive term that simulates the effect of the Pauli principle—the orthogonalization of the conduction-electron wave-

function to the core-electron wavefunction. The near cancellation of these two terms in the core region makes the pseudopotential weak in general, so that—to a good approximation—one can incorporate it into the electron-gas theory by perturbation methods. This procedure yields reasonable qualitative results for the "simple" metals—those in the upper left-hand corner of the periodic table—and for "nearly-free-electron" semiconductors, namely silicon and germanium.

The density-functional theory represents one further attempt to carry out the next stage of integrating the lattice potential with the mutual interaction of electrons. Here a solid is regarded as a system of many electronics with the same mutual interaction, embedded in the lattice of nuclei. The ground-state electron density distribution plays an important role, as it did in a forerunner of this theory, the Thomas–Fermi method, which we will discuss later.

The density-functional theory yields two results, one important conceptually and the other a framework for practical approximations:

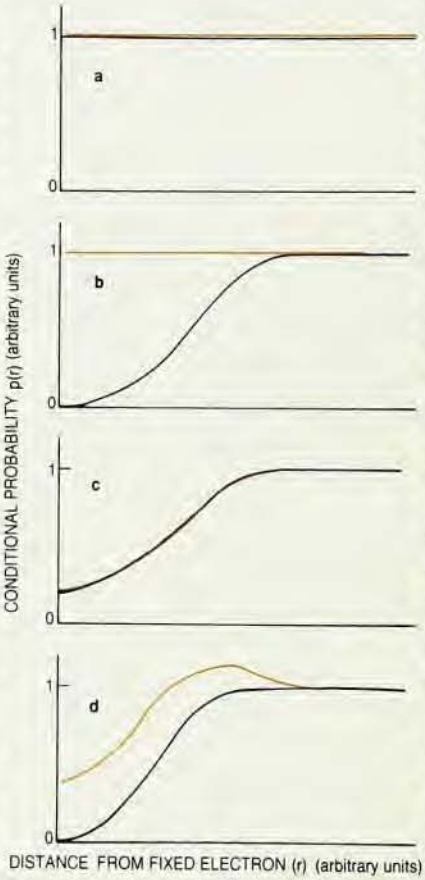
▶ The solution of the many-body ground state is reduced exactly to the solution for the ground-state density distribution given by a *one-particle* Schrödinger equation. The effective potential in the Schrödinger equation includes, in principle, all the interaction effects: the Hartree potential (the Coulomb potential due to the charge distribution when the electrons are treated as fixed), exchange (the potential due to the interaction described by the Pauli exclusion principle), and correlation (the potential due to the effect of a given electron on the overall charge distribution).

▶ An approximation for the effective potential is given by regarding a small neighborhood of the electron system as behaving like a jellium at the local density. This is a concept from the Thomas–Fermi method.

These results form the so-called "local density-functional" approximation, which, through the efforts of many researchers over the last one and one-half decades, has become a useful technique for computation of the properties of solids. In what follows, we give a qualitative account of the density functional theory, survey its many applications and examine its shortcomings in anticipation of future work.

The formalism

The whole theory is based on the remarkable theorem[2] of Pierre Hohenberg and Walter Kohn: given a mutual interaction—the Coulomb interaction in our case—the external potential $v(r)$, and hence all properties of the many-particle system, are determined by the ground-state electron density distribution $n(r)$. (The caveat that an arbitrary constant added to $v(r)$ cannot be determined is understood.) In Schrödinger's theory the converse is valid: We accept readily that $n(r)$ and all properties of the system are functionals of $v(r)$, and can be determined from $v(r)$, in principle, by the many-particle Schrödinger equation. The validity of Hohenberg and Kohn's converse is not so obvious. Indeed, the theorem was at first greeted with great skepticism, based on the feeling that the properties of the system could be determined from the one-particle density matrix and that the density $n(r)$, being the diagonal part of the density matrix, contained much less information. The proof of the Hohenberg–Kohn theorem, based on a *reductio ad absurdum* argument using the variational principle, was sufficiently simple as to receive little criticism. Instead, skeptics made valiant and ingenious attempts to construct counterexamples, from which we learned to state the theorem with more precision. The theorem has apparently so far stood the test of time.



The electron–electron interaction in an N-electron system is illustrated by these conditional probability distributions $p(r)$ of $N-1$ electrons around one electron with given spin, situated at $r=0$. In the Hartree approximation (**a**) all electrons are treated as independent and $p(r)$ is structureless. In the Hartree–Fock approximation (**b**) the many-electron wavefunction reflects the Pauli exclusion principle. Thus around the electron at $r=0$ (and around all other electrons as well) we see the so-called exchange hole, in which the density of electrons having like spin is reduced (black curve) while the density of electrons having opposite spin is unaffected (colored curve). The underlying thought of the local-density approximation (**c**) is that electrons of both spins show the same hole, the exchange-correlation hole. The local-spin-density approximation (**d**) allows for differently shaped holes in the distributions of electrons with equal spin (black curve) and electrons with opposite spin (colored curve). Figure 2

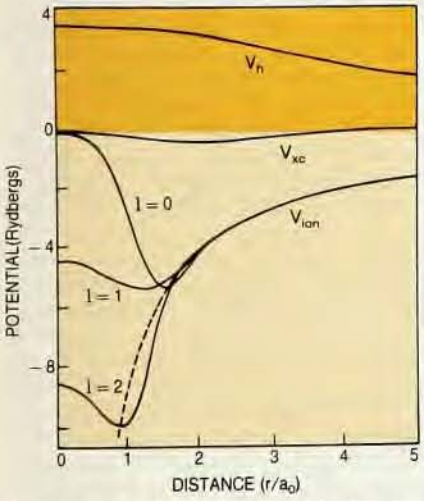
One can write the ground-state energy formally as the potential energy of the density $n(r)$ in the external potential $v(r)$ plus a term $F[n]$ which, because of the theorem, is a functional of $n(r)$ only. (A functional is a quantity that is dependent on a variable function and not on a number of discrete independent variables.) A variational theorem is then derived, which states that, given the external potential $v(r)$, the correct electron density is the one that yields the smallest value for this expression of the energy. The variational equation in principle determines the density, given $v(r)$. The problem is that the functional dependence of the remainder $F[n]$, which contains the kinetic, Coulomb, exchange and correlation energies, is not really known. If the local-density approximation as described above is applied to the entire $F[n]$ term, we get the Thomas–Fermi equation together with the correction due to exchange and correlation.

Kohn and Sham[3] noted that in a non-interacting system, $F[n]$ is the single-particle kinetic energy and the solution of the variational equation in functional derivatives of $F[n]$ is equivalent to the solution of the one-particle Schrödinger equation for the density. In an interacting system, on the other hand, the total energy may be separated into the kinetic energy of the non-interacting system at the same density distribution, plus the rest, which includes the lattice potential energy, corrections to the kinetic energy, the Hartree potential, exchange and correlation. We may then regard solving the equation for the functional derivative as equivalent to solving a one-particle Schrödinger equation with an effective potential that is just the functional derivative of the total energy minus the kinetic energy of the corresponding non-interacting system. In this sense, the solution of the many-electron ground state reduces to that of a one-electron Schrödinger equation. All the interaction physics is lumped into the construction of the effective potential.

This reduction becomes a practical scheme when we are given a reasonable approximation for the effective poten-

tial. In the spirit of two of its antecedents (the Thomas–Fermi approximation and the Slater exchange), the local-density approximation is applied to the total energy term minus the kinetic energy of the non-interacting system of equivalent density distribution. That is, the energy term is given by replacing it in a locality by that of the homogeneous electron gas at the same density. Because the rapidly varying kinetic energy is not approximated by the local expression as in the Thomas–Fermi method, one can retain the atomic shell structure in the density distribution and avoid the divergence of the density at the nuclear potential singularity.

The exchange and correlation part of the effective potential is just the corresponding part of the chemical potential of the homogeneous electron gas at the local density. The exchange term alone turns out to be the Dirac expression for the exchange term in the Thomas–Fermi approximation, an expression that is two-thirds of the Slater exchange; in the Slater exchange the energy changes as the cube root of the electron density $n(r)$. Historically, this led to two approaches to the effective potential:



Components of the atomic potential for silicon, as given by the self-consistent solution in the local-density functional method. The ion core is described by a pseudopotential ion that at large distances behaves like an attractive Coulomb potential ($-Z_v/r$, broken line) arising from the silicon valence charge. At short distances the pseudopotential depends on angular momentum to describe correctly the influence of the atomic core. The repulsive potential V_n, due to the electrostatic interaction between valence electrons, is cancelled at large distances by the attractive ion-core potential V_{ion}. Because of this cancellation, the integrated energy of the local-density functional theory's exchange and correlation potential V_{xc} amounts to about 30% of the total energy of the silicon valence system. The Rydberg is 13.6 electron volts; a_0, the Bohr radius, is 0.529 angstrom. Figure 3

▶ What might be termed the chemical approach was initiated by John C. Slater's suggestion of incorporating into his exchange a multiplicative constant α that is determined phenomenologically, for example by reproducing the ground-state energy of an atom. This is the widely used "Xα method."

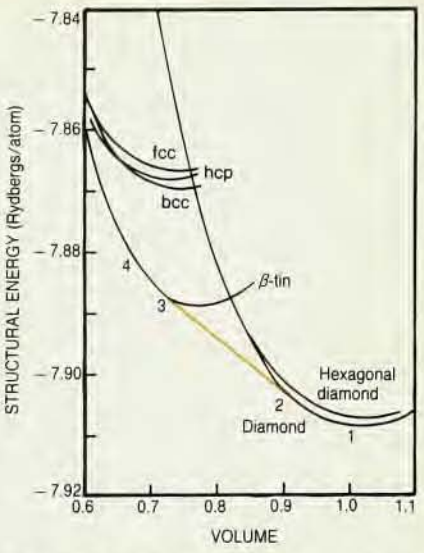
▶ What might be termed the physical approach is the original local-density approximation of Kohn and Sham, which depends on a knowledge of exchange and correlation in the jellium. Researchers have put considerable effort into deriving convenient correlation functionals interpolating between regimes of high and low density.

So far we have obtained only the ground-state energy and density. Extension to finite temperatures is straightforward by virtue of David Mermin's proof[4] of the Hohenberg–Kohn theorem for the free energy at a finite temperature. Extension to finite magnetic fields is possible by enlarging the density variable to include unequal local spin densities; this is known as the local spin-density approximation.

When we look at some results we will see how accurately the local-density approximation yields ground-state properties. Some researchers have developed an argument that gives a measure of understanding to the success of the local-density approximation. One can express the exchange- and correlation-energy functional exactly in terms of a pair-correlation function, which may be regarded as describing the exchange–correlation hole around an electron. The local-density approximation as described here amounts to using the exchange–correlation hole in a homogeneous electron gas at the local density. It preserves the sum rules that express the total charge of the exchange–correlation hole as being equivalent to the removal of one electron (see figure 2).

Ground-state silicon

Most of the calculations that apply the local-density functional approximation to atoms, molecules and solids involve complicated numerical manipulations that are necessary to solve Schrödinger's equation, and therefore the accuracy of the numerical solution for a given density functional is often unknown. Obviously, the risk of numerical inaccuracy increases in going from simple systems such as atoms to complicated systems such as semiconductors or transition metal surfaces. Among the variety of high-quality calculations, we select as an illustrative example recent work by a group at the University of California, Berkeley, on the phase transformation and lattice dynamics of silicon.[5] Later, we will give a brief survey of other applications.



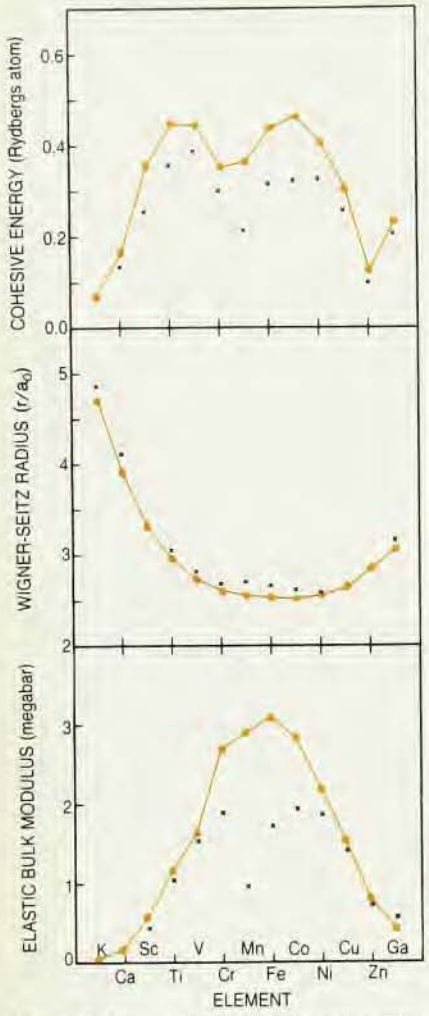
Calculated total structural energy of silicon atoms held to each of six hypothetical lattice arrangements. Energies are a function of atomic volume, which is normalized here to the actual ground-state volume of silicon in its natural diamond structure. Under external pressure, silicon follows the path 1–2–3–4 and undergoes a phase change to the β-tin structure. At pressures corresponding to volumes between points 2 and 3 silicon is a mixture of the two structures. The tangent points 2 and 3 give the calculated phase transition volumes, which are in excellent agreement with experiment. (From reference 5.) Figure 4

Other than using the local-density functional approximation to describe exchange and correlation, the silicon calculations employ two more approximations:

▶ the "frozen core" approximation, which considers nuclei plus core electrons as a neutral "frozen" unit that is not responsive to changes in chemical environment, and

▶ the simulation of this core by an ion-core pseudopotential.

The pseudopotential concept *per se* is very general and has been used extensively in the past with various degrees of sophistication (see Marvin L. Cohen's review in PHYSICS TODAY[6]). Recently, however, several groups have developed theoretical schemes to produce parameter-free pseudopotentials of high accuracy. By construction, these potentials exactly simulate the interaction of the valence electrons with the cores in the atomic limit (that is, for a chosen atomic reference state). And, also by construction, they guarantee maximum transferability to other chemical environments. The important feature is that only the ion-cores are represented by pseudopotentials while the valence electron screening part is recalculated in the actual chemical situation. This is in contrast to the empirical one-electron pseudopo-



Ground-state properties of some transition metals, as calculated (curves) and measured (unconnected points). The Wigner–Seitz radius is a measure of the lattice constant. These graphs focus on the elements for which the local-density functional calculations are least accurate: the middle of the periodic table's first row of transition metals, in which magnetic anomalies occur that are not described in the chosen spin-unresolved formalism. (From reference 8.) Figure 5

tentials we mentioned earlier. Figure 3 illustrates the advantage of local-density functional calculations over other many-electron schemes. Few smooth potentials describe systems of atoms, molecules or solids to a high degree of accuracy. Transition metal atoms can be treated with the same accuracy and rigor as can simple s or p atoms. While not being weak perturbations, these ion-core pseudopotentials have no sharp singularities (in contrast to the "true" ion cores) and accurate self-consistent numerical solutions of Schrödinger's equation are generally feasible using expansion techniques.

The results of the silicon calculations are impressive. Among the six crystal structures that cover about 90% of all existing elemental solids (diamond, hexagonal diamond, β-tin, hcp, bcc and fcc), the calculations show correctly that the diamond structure is the most stable when the temperature is zero and there is no external pressure. The stable diamond structure is separated by only about 20 millielectron volts per atom from the hexagonal modification. This small energy arises from structural differences beyond the third nearest neighbors, which, for example, separate compound-semiconductors into zinc-blende and wurtzite structures. In the diamond structure, calculated and measured static ground-state properties, such as the equilibrium lattice constant, bulk modulus and cohesive energy, agree to better than a few percent. The calculations also show correctly that silicon will transform to the β-tin structure under high pressure. The transition volumes, which are extremely sensitive quantities, are reproduced to within 1% of experimental values (see figure 4).

One can calculate the effects of lattice dynamics by assuming adiabatic behavior and comparing crystal energy differences for various "frozen-in" phonon distortions. Calculations match measurements to within better than 3% for acoustic and optical phonons on the Brillouin zone surface and at the zone center. In particular, calculations of the very soft transverse acoustic mode, characteristic of shear instabilities in tetrahedral semiconductors, show only a 1% error. Moreover, anharmonic effects are described within 10%, which is remarkable in view of the fact that many theories cannot even reproduce their correct signs.

Overall, we should regard the silicon results as an important confirmation of the validity of the local-density functional approach for studying electronic ground-state properties of systems with s and p valence electrons and with sizeable external lattice potential. All these results stand in strong contrast to the poor quality of calculated excitation energies. If one takes the energy eigenvalues of the local-density functional Schrödinger equation to be excitation energies, just as Koopmans's theorem tells us to do in Hartree–Fock theory, errors up to about 50% occur in comparison to experimental gap values. We will come back to this dilemma after looking at some other applications.

Other applications

The example of silicon illustrates the great accuracy possible using the local-density functional approximation. The same approximation has been applied to a wide range of systems; the number of contributions is enormous and certainly is not exhausted in the following review.

Atoms. Because many atoms have unpaired electrons, the two-component spin-up and spin-down density distributions replace the single density variable. Calculations have been done for many atoms. While the computations are much simpler than in the unrestricted Hartree–Fock scheme, many of the calculated atomic properties are at least as accurate. The list includes total energy, ionization potential (as the difference between the total energies of the ion and the atom), approximate multiplet energies and hyperfine interaction. Trends in the properties of elements in the periodic table, reflecting Hund's rule on the order of filling of orbitals, are reproduced and the correct ground states are generally predicted. There are a few exceptions, such as titanium, vanadium and cobalt, where the $d^{n-1}s^1$ configuration is favored by local-spin-density calculations over the correct $d^{n-2}s^2$ groundstate.

Molecules. A major difficulty in molecular calculations is still accuracy in the solution of the Schrödinger equation. Investigators have taken advantage of the often-used scattered-wave method[7] to study molecules as large as tetrathiofulvalene–tetracyanoquinodimethane (known as TTF-TCNQ). Several other techniques have been used since with considerable success. Other researchers have made quantitative studies of many diatomic molecules using the "linear combination of muffin-tin orbitals" method and the "linear combination of Gaussian orbitals" method. Generally, experimental trends in molecular binding energy, bondlength and vibration frequency are reproduced with less than 10% error. The results of the local-spin-density approximation are better than those of the Hartree–Fock approximation because of the former's inclusion of correlation effects. Particularly striking results have been obtained by researchers in Jülich on the reversal of the dipole moments in CO and NO and the binding of the group IIA dimers, Be_2 through Ba_2, which are reproduced by the local-spin-density method but not by the Hartree–Fock scheme.

Bulk metals. The most comprehensive description of metal properties is undoubtedly given in the book by Victor L. Moruzzi, James F. Janak and Arthur R. Williams.[8] They report properties for the first 50 elemental metals, calculated with a band structure scheme (Kohn–Korringa–Rostocker) using the local-density functional or the local-spin-density approximation for exchange and correlation. The agreement with experiment of binding energy, lattice parameter and bulk modulus is generally around 10%, except where certain magnetic anomalies occur, as shown in figure 5.

There has been a most exciting advance in the theory of magnetism. The

works of many researchers demonstrate the theoretical ability of the local-spin-density approximation to predict ferromagnetism and to calculate the magnetic moments and susceptibilities. The splitting of the exchange interaction between spin-up and spin-down bands seems to be in good agreement with photoemission data for iron and cobalt but not nickel. Again this points to some inadequacy of interpreting the method's eigenvalues as excitation energies.

Another application that is being explored is the equation of state of metals under high pressure. There is evidence of improvement over approximations of the Thomas–Fermi type.

Bulk semiconductors and insulators. Semiconductors and insulators exhibit open (not closely packed) structures and strong lattice potentials. This necessitates very accurate numerical procedures to solve Schrödinger's equation. Investigators have today developed pseudopotential methods, local-orbital methods and methods similar to the augmented-plane-wave scheme that yield highly accurate numerical solutions. They apply these techniques to the study of systems ranging from the usual semiconductors to large-gap insulators such as diamond, boron nitride, lithium fluoride and the rare-gas solids. For calculating cohesive energies and other ground state properties the local-density functional approximation is a vast improvement over the more complicated Hartree–Fock approximation, particularly in the rare-gas solids where the latter method gives *no* binding. The calculated gap energies, however, are much too small—a familiar failure of the local-density functional approach.

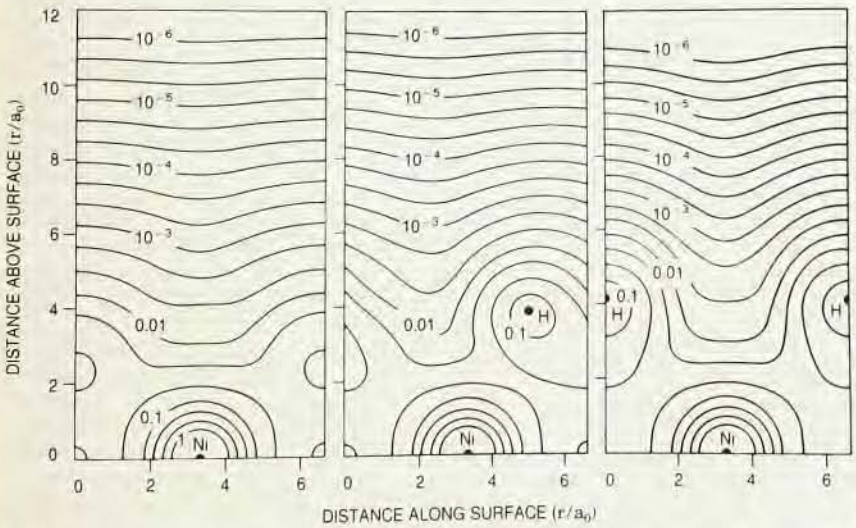
Defects. Localized defects or impurities in either metals or semiconductors are challenging systems for the theorist. Of particular interest are systems of intermediate defect-localization and strong coupling to the host environment. These systems have been most widely described by the same approach used for large molecules or clusters. Among the many systems that researchers have looked into, we will mention the Xα-scattered wave calculations used to describe lattice defects and transition-metal impurities in semiconductors. In spite of the more qualitative insight gained by these calculations it is generally felt that prohibitively large cluster sizes are necessary to obtain quantitative results.

One can circumvent these difficulties by using a Green's function approach to the isolated defect in an otherwise perfect host crystal. For instance, investigators have used this method to study the electronic structure of magnetic impurities in metals such as copper and silver. Using a spin-resolved density functional, they find the calculated magnetic moments in very good agreement with room-temperature susceptibility measurements. Local-density-of-states calculations are in general agreement with the Anderson model, which predicts well-resolved resonances of local moments in metals.

A series of self-consistent local-density functional calculations using Green's functions has recently also been applied to semiconductors. The calculations concern vacancies in silicon, diamond, gallium arsenide, gallium phosphide and also substitutional impurities such as hydrogen, carbon, nitrogen or oxygen (see figure 1). All the defects give rise to localized states in the bandgap of the host semiconductor. Electronic screening effects, described self-consistently in the local-density functional framework, are particularly important in the case of the more-ionic hosts such as the gallium compounds mentioned above, and produce defect potentials that are highly localized (to within the nearest-neighbor range). The calculations show that lattice relaxation effects around the defect are of equal or even higher importance. As Gene Baraff and his collaborators demonstrated, in the case of the silicon vacancy the distortions are strongly affected by the amount of charge localized at the defect. For this particular example, they calculated that (Jahn–Teller) coupling to the lattice produced large, energy-lowering distortions, strong enough to overcome the Coloumb repulsion of extra electrons in the localized state. Thus they predicted an "inverted-level" scheme or "negative U" situation for the simple defect and this was later verified experimentally.

Surfaces. The use of local-density functional approximations in describing metal or semiconductor surfaces began with the pioneering work of Norton Lang and Kohn on structureless "jellium" surface models. While general features like (Friedel) charge oscillations were found, the absence of discrete atomic structure in the model led to sizeable errors in properties such as work functions or surface energies, errors that could be corrected to some extent by second-order perturbation theory of the psuedopotential.

At Bell laboratories, Joel Appelbaum and Donald Hamann introduced calculations with real surface structures in the early 1970s and this has since led to a wealth of work on both metal and semiconductor surfaces. These calculations involve rather complicated techniques that are necessary to solve Schrödinger's equation in the surface region. Because of the lack of numerical accuracy, earlier calculations were mainly aimed at conveying global surface characteristics and general spectral features, such as the position and dispersion of surface states, which could be compared qualitatively with data from photoemission surface experiments. This "finger-printing" technique proved to be particularly powerful in determining the structure of adsorbates and the nature of surface reconstructions. These calculations were complemented by local-density functional calculations of states in the higher energy range (10–100 eV), in-



Surface charge density for nickel as given by local-density functional calculations. Contours of equal charge density are logarithmically spaced to emphasize the exponential decay of charge in the vacuum above the metal's surface. The figure shows clean nickel (left) and hydrogen-covered nickel (middle and right) with the hydrogen in slightly different positions. The small changes in the 10^{-5} regime, which is about 5 angstroms above the surface, can be measured by diffracting slow helium atoms off the surface. Distances are given in units of the Bohr radius a_0; charge density is given in electrons $/a_0^3$. (From reference 10.) Figure 6

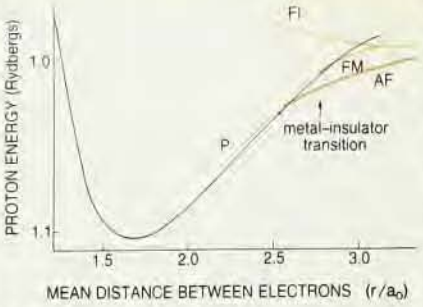
tended to explain the results of low-energy-electron diffraction experiments.

Improved numerical techniques are leading to more quantitative results, comparable to results of the studies on bulk metals and semiconductors. Recent calculations on the (111) surface of aluminum, for instance, yield an excellent value for the work function (about 1% error). Groups at Bell, IBM and Berkeley have had encouraging results in their attempts to determine structure by the direct minimization of the surface energy. Hamann demonstrated recently that local-density functional calculations on surfaces can be refined to a degree that permits the determination of the small, structure-induced charge modulations that occur some 5 to 10 Å outside the surface. These charge-density modulations, which are as small as 10^{-5} times the charge density in the chemical bonds in the crystal, can be determined experimentally by helium diffraction. Some examples are shown in figure 6.

Electrons in inversion layers. The local-density functional approach has also been used to study quasi-two-dimensional systems of electrons. One can form these systems at the insulator–semiconductor interface in a metal–insulator–semiconductor field-effect transistor (MISFET) by applying an electric field normal to the sandwich. We can account for the variation of the effective potential normal to the interface, within the density-functional formalism. The interaction effect of this many-electron system appears well given by a combination of the local-density functional approximation and the Hamiltonian of an effective-mass-like model, as recent calculations of inversion-layer energy bands demonstrate.

In silicion or germanium, the electrons have both spin degeneracy and valley degeneracy, with the latter being due to the existence of more than one conduction band minimum. Some researchers have suggested that the low-density behavior of the inversion layer may be due to a phase transition involving the removal of the valley or spin degeneracies. Measurements of conductivity oscillation in high magnetic fields and precise measurements of optical absorption have since confirmed the existence of such a phase transition. This is an interesting two-dimensional analog of a ferromagnetic gas, which we will describe next.

Metal-insulator transitions. Several groups have recently used the local-spin-density approximation to study the occurrence of metal–insulator transitions in a three-dimensional gas of interacting electrons. The problem is simulated by Nevill Mott's model of electrons in a lattice of fixed point charges: a hydrogen lattice. The ground state is determined variationally as a function of density. If the high space-group symmetry of one atom per unit cell is maintained, the overall picture shows two phase transitions: a second-order transition at higher density from the paramagnetic metal to the spin-ordered metal, followed by a first-order transition at lower density to a spin-ordered insulator. If the symmetry is allowed to be lowered to two atoms per unit cell, these two transitions are replaced by a single transition into an *anti*ferromagnetic ground state. These findings demonstrate the ability of the local-spin-density approximation to explain the occurance of phase transitions as the result of a trade-off between exchange–correlation energy and kinetic energy (see figure 7).



Total energy per proton as a function of density (or the mean distance between electrons), calculated for a system of hydrogen atoms on a hypothetical body-centered-cubic lattice. The system undergoes various phase transitions from the paramagnetic phase (*P*) as the density is lowered. *FM* is the ferromagnetic, metallic phase; *FI* is the ferromagnetic, insulating phase, and *AF* is the antiferromagnetic phase, which shows a metal–insulator transition at the position indicated by the arrow. (From reference 11.) Figure 7

Future directions

We have pointed out many successful applications of local-density functional or local-spin-density theory. However, there are also failures, notably in two areas:

- the ground states of some open-shell systems, such as 3d transition-element atoms, that are sensitive to the exchange and correlation energy, and
- excited "quasiparticle" states of the system.

With respect to the first problem, there is much interest in working on improved functionals that go beyond the original Thomas–Fermi-like, strictly local-density approach, to include some nonlocal effects. The new approaches focus on the exact fulfillment of the corrrelation sum rule we mentioned earlier, and on the exact elimination of spurious terms left over by the incomplete cancellation of the electron self-interaction in the Hartree term. However, functionals including the nonlocality of real space may still need further improvement. Recent studies of atomic multiplets seem to indicate the need for the nonlocal functional to depend on the particular *symmetry* of the state under consideration.

The description of excitations using the local-density functional scheme is a further major remaining problem. By analogy to Koopmans's theorem in Hartree–Fock theory, the Lagrange parameters or local-density functional eigenvalues are often interpreted as approximate excitation energies. This can lead to sizeable errors—for example, atomic ionization energies that are underestimated by several volts, or band gaps that are up to 50% too small. The error originates from

- the neglect of electronic relaxation upon excitation (meaning that Koopmans's theorem is not valid), and
- the "misinterpretation" of eigenvalues.

In local-density functional schemes the eigenvalues are equivalent to the *derivative* of the total energy with respect to occupation, evaluated at a particular configuration, such as the ground state. The eigenvalues are thus not finite differences of total energies as in Hartree–Fock theory. Slater was the first of many to recognize that the exact elimination of electron self-interaction (as is done in Hartree–Fock but is not necessary in a local-density functional theory) would account for a large fraction of the error in excitation energies. For localized systems one can eliminate most of the problem by considering finite differences in the total (local-density functional) energy either directly, by comparing two calculations, or by using Slater's transition-state arugment. Little is known, however, about what to do for extended systems, where it may be inherently necessary to consider explicitly the energy dependence of the self-energy operator in describing excitations. Kohn and Sham have pointed out that in a formal sense the self-energy is also determined by the ground-state density, and they have given an approximate scheme for calculating it in the limit of low-energy excitations.

In the framework of many-body perturbation theory on the other hand, we can determine the self-energy iteratively from the complete excitation spectrum. Investigators have undertaken several attempts along these lines to calculate quasiparticle excitation spectra. While these calculations clearly go beyond any local-density functional theory, the hope is to arrive at schemes that allow us to calculate excitation energies with a simplicity comparable to the very successful local-density functional ground-state calculations.

If we accept the evidence gained so

far, then the density functional method, in conjunction with practical approximations, is capable of describing ground states of interacting systems. It is, therefore, a useful tool for further exploration in solid-state physics. Used in combination with low-energy-electron diffraction or surface-atom scattering, we can investigate the fascinating variety of atomic arrangements on surfaces and their underlying physical causes. In conjuntion with x-ray scattering and x-ray absorption, we can deduce structures and bonding properties of new, complex materials. Calculations of solid structures subject to various external factors such as pressure, can help us decipher structural phase transitions. Accurate calculations of the effective interaction between atoms in solids or on surfaces may become indispensable to new methods of statistical mechanics that are being applied to the study of properties associated with phase transitions.

Simple models, which are often invaluable as the cutting edge in understanding new phenomena, sometimes contain assumptions that we can check by such precise calculations, and conversely, precise calculations can point the way for the construction of simple models. Although efforts to perfect the density functional construction might lead to the making of an automaton that could calculate certain properties from scanty input, we do not regard this as a valid end in itself. Such computations would be limited by the physical ideas built into the theory. However, the interplay between observation, physical insight and the results of such careful computations will, we hope, lead to new vistas in the physics of the condensed state.

References

1. J. C. Phillips, L. Kleinman, Phys. Rev. **116**, 287 (1959).
2. P. Hohenberg, W. Kohn, Phys. Rev. **136**, B864 (1964).
3. W. Kohn, L. J. Sham, Phys. Rev. **140**, A1133 (1965).
4. N. D. Mermin, Phys. Rev. **137A**, 1441 (1965).
5. M. T. Yin, M. L. Cohen, Phys. Rev. Lett. **45**, 1004 (1980).
6. M. L. Cohen, PHYSICS TODAY, July 1979, page 40.
7. K. Johnson, J. C. Slater, PHYSICS TODAY, Oct. 1974, page 34.
8. V. L. Moruzzi, J. F. Janak, A. R. Williams, *Calculated Electronic Properties of Metals*, Pergamon, New York (1978).
9. G. B. Bachelet, G. A. Baraff, M. Schlüter, Phys. Rev. B **24**, 4736 (1981); G. B. Bachelet, H. Greenside, G. A. Baraff, M. Schlüter, Phys. Rev. B **24**, 4745 (1981).
10. D. R. Hamann, Phys. Rev. Lett. **46**, 1227 (1981).
11. P. Kelly, O. K. Andersen, T. M. Rice, to be published in Phys. Rev. □