PHYSICAL REVIEW B VOLUME 39, NUMBER 17 15 JUNE 1989-I

Tight-binding models and density-functional theory

W. Matthew C. Foulkes*
Theory of Condensed Matter Group, Cavendish Laboratory, University of Cambridge,
Madingley Road, Cambridge CB3 0HE, United Kingdom

Roger Haydock†
Department of Physics and Materials Science Institute, University of Oregon, Eugene, Oregon 97403-1274

(Received 20 October 1988)

Semiempirical tight-binding models have been widely used but the details of their relationship to more fundamental theories have never been clear and so they have usually been treated as fitting and interpolation schemes rather than as quantitative calculational tools. In this paper, we show how simple tight-binding models can be understood as stationary approximations to self-consistent density-functional theory and give prescriptions for calculating all the required potentials and matrix elements. Some preliminary applications have given encouraging results.

INTRODUCTION

One of the central problems of theoretical solid-state physics is to find ways of calculating the forces acting between the atoms in a solid. This is particularly important near defects such as vacancies, dislocations, and grain boundaries, since the interatomic forces affect the way defects behave in response to externally applied stresses and changes in temperature and hence determine many of the macroscopic mechanical and thermal properties of the solid. Despite their importance, however, interatomic forces are hard to calculate. This paper describes an attempt to improve one of the available methods, the tight-binding method.

Methods for calculating interatomic forces may be divided, roughly, into two classes: those which attempt a proper quantum-mechanical treatment of the electrons and those which do not. Into the latter category fall most pair-potential methods (an exception, perhaps, should be made for the volume-dependent pair-potential description of weak pseudopotential solids[1] derived from second-order pseudopotential perturbation theory) and most methods using three- or higher-body interatomic potentials (exceptions here might be made for some embedded-atom methods[2]). Among the quantum-mechanical approaches, the most accurate are based on density-functional theory[3] and include the standard band-structure methods such as the linear muffin-tin orbitals (LMTO) method, the linear augmented-plane-wave (LAPW) method, and the use of plane-wave basis sets with norm-conserving pseudopotentials. Less accurate approaches include the embedded-atom and pseudopotential perturbation methods already mentioned, the use of the Gordon-Kim model,[4] and the tight-binding method which is the principal subject of this paper.

Although the methods which are not based on quantum mechanics are adequate in many cases, there is considerable evidence that it is far from easy to find reliable classical interatomic potentials for many important classes of solids. An instructive example is provided by the work of Biswas and Hamann[5] on silicon. They attempted to fit a three-body potential of a fairly general form to a large number of properties of silicon in many different crystal structures (most of these were purely hypothetical and so the required properties were obtained from accurate self-consistent density-functional calculations). The database was about the largest of its kind that could have been collected, and the numerical work and fitting procedures seem to have been of a high quality; but nevertheless, as soon as the potentials derived were used to study defects which had local arrangements of atoms markedly different from any included in the database, they were found to give unreliable and inaccurate answers. The lack of "transferability" was particularly pronounced when looking at highly undercoordinated structures. This is a clear demonstration that the physical laws governing the interatomic forces in silicon are not well represented by simple interatomic potentials and suggests that more sophisticated approaches, presumably based on quantum theory, are needed.

Among the quantum-mechanical approaches, only full self-consistent density-functional calculations with a good basis set seem capable of giving interatomic forces which can be trusted quantitatively in most solids. However, most self-consistent density-functional calculations are cast in reciprocal space and so rely on the periodicity of the arrangement of atoms in the solid. They are inapplicable when this periodicity is disrupted, as happens near a defect (point, linear, or planar) or in a glass. For defect calculations, the periodicity can be restored by the artificial device of considering a periodic array of identical defects, but to ensure that the defects in such an array are far enough apart to be treated independently it is often necessary to use very large unit cells. Although great improvements have recently been made[6] in the sizes of unit cells which can be treated, such "supercell" calculations are still time consuming, computationally intensive, and rather limited as to the sort of questions they can answer. There is no doubt that supercell calculations

©1989 The American Physical Society

provide an invaluable benchmark and are becoming rapidly more flexible but there is still room for simpler and less accurate approaches, both because they can deal with more complicated arrangements of atoms and because they are often easier to interpret and provide more physical insight.

Among the less accurate quantum-mechanical methods, the semiempirical tight-binding (SETB) method is probably the most widely used and generally useful. Its successes have been many and varied,[7] but the method is far from reliable. There is no doubt that SETB calculations are easy to do and that the results, expressed in terms of quantities such as local densities of states and chemical bond orders, are easy to interpret and give a very physical description of what is going on in the solid. However, since even simple qualitative results are not always reliable, it is often very difficult to know whether to trust SETB calculations or not. A couple of examples will illustrate both the strengths and the weaknesses of the SETB approach.

An example of a successful SETB calculation is the work in which Chadi[8] predicted the buckling (asymmetric dimerization) of the silicon (100) surface. Subsequent self-consistent density-functional calculations[9] disagreed about whether this reconstruction occurred or not, but more recent experimental and theoretical work[10] on both Si and Ge has tilted the balance back in favor of the original SETB results.

Examples of some of the problems which can arise when using SETB are provided by the work of Hashimoto[11] and of Sayers,[12] who both looked at the problem of impurity-induced intergranular embrittlement in metals. Both came to the conclusion that embrittlement resulted from a decohesion mechanism. According to Sayers, the weak bonds are those between the impurity atoms at the grain boundary and the neighboring host metal atoms (first-layer decohesion), whereas Hashimoto attributes the embrittlement to a weakening of the bonds between pairs of host metal atoms which happen to lie near the impurity-contaminated grain boundary (second-layer decohesion). However, a recent self-consistent pseudopotential calculation[13] seems to rule out both these conclusions, and Goodwin, following an earlier suggestion of Haydock,[14] believes that embrittlement arises not because of a decrease in cohesion at all, but because the impurities like to form covalent and directional bonds which are not easily broken and remade. This means that atoms at the boundary are less able than those in the bulk to dissipate local concentrations of stress by ductile flow (so blunting a crack tip, for example) and so the boundary becomes more prone to brittle fracture.

The problems exemplified in this work arise, at least in part, from the excessive flexibility of SETB models as fitting schemes. In a typical tight-binding calculation there are between five and ten unknown functions of the interatomic distances which must be fitted to the results of experiments or to more accurate calculations. It is not surprising, therefore, that there is usually little difficulty in finding parameters which reproduce the chosen data with reasonable accuracy. Unfortunately, just as happened in the case of the Biswas-Hamann potential for silicon, this does not ensure that the derived parameters are physical or transferable. This can only be guaranteed if the meanings of the parameters and the relationships between them are understood in more detail, and if ways are found of calculating them directly using a more fundamental theory. The most obvious choice of fundamental theory is density-functional theory.

The relationship between SETB theory and density-functional theory is straightforward enough on a rough intuitive level, but the details are only just beginning to become clear.[15,16] The sorts of questions which must be answered are the following. Why are most tight-binding calculations non-self-consistent? How do they include the double-counting energy present in self-consistent calculations? Can one sensibly assume that the tight-binding basis functions are orthogonal? Is it reasonable to take the Hamiltonian matrix elements to be simple pairwise functions of the interatomic distances? And can one ignore all crystal-field terms as is usually assumed?

This paper is intended as the first of several in which we hope to answer some of these questions and to show how to derive reliable and transferable tight-binding parameters from density-functional theory. These can then be applied, in conjunction with the recursion method, to study the behavior of complicated defects in real solids. All the calculations can be cast in real space[7,17] and so this approach avoids the need to impose an artificial periodicity by using supercells. Many real-space tight-binding defect calculations have already been done, of course, but they have almost all suffered from uncertainty about whether the parameters involved are sensible and trustworthy or not; the aim of our work is to help remove this uncertainty.

The rest of this paper is organized as follows. In Sec. I we describe and compare the density-functional and semiempirical tight-binding approaches to calculating total energies in solids. We show how the questions which must be addressed in attempting to relate the two fall naturally into two groups: questions concerning the form of the tight-binding expression for the total energy; and questions concerning the form of the tight-binding secular equation. We restrict ourselves to considering questions in the first of these groups for the remainder of the paper. In Sec. II, we derive a rather unusual stationary principle in density-functional theory, and in Sec. III show how it leads to an expression for the total energy of the form assumed in simple tight-binding models. This is tested on some simple examples in Sec. IV. Finally, we conclude by reiterating what has been accomplished and discussing what is still to be done.

I. DENSITY-FUNCTIONAL THEORY AND SEMIEMPIRICAL TIGHT-BINDING MODELS

We start this section with a brief summary of some parts of Hohenberg-Kohn-Sham density-functional theory (DFT).[3] Then we describe what we mean by a semiempirical tight-binding (SETB) model and discuss the points of similarity and difference between SETB model calculations and full self-consistent DFT calculations.

A. Density-functional theory

Once the Born-Oppenheimer approximation has been made, the ground-state electronic energy of a solid (for any given arrangement of the nuclei) can be obtained by solving the Schrödinger equation for the interacting electrons moving in the fixed external potential due to the frozen nuclei. According to DFT, this ground-state energy is the minimum value of a functional, $E[n(r)]$, of the electron density, $n(r)$. The functional is defined for all normalized densities $n(r)$, but takes its minimum value when the density is the ground-state density $n_0(r)$.

It is useful to divide the total-energy functional into two parts,

$$E[n]=T_s[n]+F[n] \ , \tag{1}$$

where $T_s[n]$ is the kinetic energy of an imaginary noninteracting (except via the Pauli exclusion principle) electron gas moving in that external potential which induces a ground-state density equal to $n(r)$. [Since not all densities are possible ground-state densities of a system of noninteracting electrons, the functionals T_s and, hence, F are not always defined (although, see Levy[18]). Densities which do correspond to possible noninteracting ground states are called "wave-function noninteracting v representable" and these are the only densities for which Eq. (1) makes sense.] $T_s[n]$ is not the same as the kinetic energy of the real interacting system, but the hope is that it is roughly similar in magnitude.

$F[n]$ must contain some simple electrostatic terms and so we write

$$F[n]=\int V_{\text{nucl}}(r)n(r)d^3r +\frac{1}{2}\int\int\frac{n(r)n(r')}{|r-r'|}d^3r'\,d^3r+E_{\text{xc}}[n(r)] \ . \tag{2}$$

$V_{\text{nucl}}(r)$ is the Coulomb potential due to the nuclear charges and this equation acts as a definition of the exchange and correlation functional, $E_{\text{xc}}[n(r)]$. If we now define the Hartree and electron-nuclear Coulomb energy functionals by

$$E_H[n(r)]=\frac{1}{2}\int\int\frac{n(r)n(r')}{|r-r'|}d^3r\,d^3r' \ ,$$

$$E_{\text{el-nucl}}[n(r)]=\int V_{\text{nucl}}(r)n(r)d^3r \ ,$$

then Eq. (2) becomes

$$F[n(r)]=E_{\text{el-nucl}}[n(r)]+E_H[n(r)]+E_{\text{xc}}[n(r)] \ . \tag{3}$$

The exact form of $E_{\text{xc}}[n]$ is not known, but it is usually well represented by a local-density approximation (LDA),

$$E_{\text{xc}}[n(r)]=\int \varepsilon_{\text{xc}}(n(r))n(r)d^3r \ , \tag{4}$$

where $\varepsilon_{\text{xc}}(n)$ is the exchange and correlation energy per electron in a uniform electron gas of density n. Good parametrized versions of $\varepsilon_{\text{xc}}(n)$ have been obtained using diagrammatic[19] and Monte Carlo[20,21] methods and these can be used in combination with Eq. (4) to provide surprisingly reliable estimates of $E_{\text{xc}}[n(r)]$ in real calculations.

So, to evaluate $E[n(r)]$ at a given density, $n(r)$, the procedure is as follows.

(a) Find the one-electron potential $V(r)$ (unique to within a constant) which gives the imaginary system of noninteracting electrons a ground-state density equal to $n(r)$. The Schrödinger equation for the noninteracting system separates to give one-electron equations of the form

$$[-\tfrac{1}{2}\nabla^2+V(r)]\psi_i(r)=\varepsilon_i\psi_i(r) \ , \tag{5}$$

and the required density, $n(r)$, is the sum of the densities associated with the lowest N (equal to the number of electrons in the system) one-electron spin eigenfunctions:

$$n(r)=\sum_{i=1}^{N}\psi_i^*(r)\psi_i(r) \ . \tag{6}$$

(b) $T_s[n(r)]$ is then given by

$$T_s[n(r)]=\sum_{i=1}^{N}\int\psi_i^*(r)(-\tfrac{1}{2}\nabla^2)\psi_i(r)d^3r \tag{7a}$$

$$=\sum_{i=1}^{N}\varepsilon_i-\int V(r)n(r)d^3r \ . \tag{7b}$$

(c) Within the LDA, it is straightforward to evaluate $F[n(r)]$ from $n(r)$ and so we get $E[n(r)]=T_s[n(r)]+F[n(r)]$ as required.

In fact, step (a) is not simple and so the first part of this procedure is usually reversed: one starts with a one-electron potential $V(r)$ and uses it together with Eqs. (5) and (6) to calculate the density at which the electronic energy functional is then evaluated.

We now know how to evaluate the functional at a given density, but to determine the ground-state energy and density we must find its minimum value. For densities $n(r)$ which make $E[n(r)]$ stationary, we have

$$\delta E[n(r)]=E[n(r)+\delta n(r)]-E[n(r)] =O([\delta n(r)]^2)$$

for all small fluctuations, $\delta n(r)$, satisfying

$$\int\delta n(r)d^3r=0 \ .$$

Consideration of the independent-electron problem used to generate $T_s[n(r)]$ shows that

$$\delta T_s[n(r)]=-\int V(r)\delta n(r)d^3r+O([\delta n(r)]^2) \ , \tag{8}$$

and hence,

$$\delta E[n(r)]=\delta T_s[n(r)]+\delta F[n(r)] =\int\left[\left.\frac{\delta F}{\delta n}\right|_{n(r)}-V(r)\right]\delta n(r)d^3r +O([\delta n(r)]^2) \ .$$

From this, it follows that $E[n(r)]$ is stationary whenever

$$\left.\frac{\delta F}{\delta n}\right|_{n(r)}=V(r)+\text{const} \ . \tag{9}$$

The value of the constant is arbitrary for a closed system and can be set to zero.

The functional derivative of $F[n]$ can be written in the form

$$\frac{\delta F}{\delta n}=V_{\text{nucl}}(r)+V_H[n(r)]+\mu_{\text{xc}}[n(r)] , \tag{10}$$

where $V_H[n(r)]$ is the Hartree potential,

$$V_H[n(r)]=\frac{\delta E_H}{\delta n(r)}=\int\frac{n(r')}{|r-r'|}d^3r' , \tag{11}$$

and $\mu_{\text{xc}}[n(r)]$, the exchange and correlation potential, is defined by

$$\mu_{\text{xc}}[n(r)]=\frac{\delta E_{\text{xc}}}{\delta n(r)} . \tag{12}$$

The condition [Eq. (9)] for a stationary point is seen to be a self-consistency condition: it relates the potential $V(r)$ in the independent-electron equation (5) to the density $n(r)$ generated by solving that equation. This is exactly as in Hartree theory and the procedure for finding self-consistent solutions is similar. Of course, finding a self-consistent solution means only that one has found a stationary point of $E[n(r)]$ and not necessarily the minimum value. In practice, however, the possibility of converging to stationary points other than the minimum causes little difficulty and one can usually assume that a self-consistent solution will give ground-state properties (or perhaps properties of the lowest-energy state of a particular symmetry). Once the self-consistent ground-state density, $n_0(r)$, and one-electron eigenfunctions, $\psi_i(r)$, have been found, the ground-state energy is given by

$$E_0=T_s[n_0]+F[n_0] \tag{13a}$$

$$=\sum_{i=1}^{N}\int\psi_i^*(r)(-\tfrac{1}{2}\nabla^2)\psi_i(r)d^3r+F[n_0(r)] \tag{13b}$$

$$=\sum_{i=1}^{N}\varepsilon_i+F[n_0(r)]-\int\frac{\delta F}{\delta n}\bigg|_{n_0(r)}n_0(r)d^3r \tag{13c}$$

$$=\sum_{i=1}^{N}\varepsilon_i-E_H[n_0(r)]+E_{\text{xc}}[n_0(r)]$$
$$-\int\mu_{\text{xc}}[n_0(r)]n_0(r)d^3r . \tag{13d}$$

B. Semiempirical tight-binding models

Semiempirical tight-binding calculations, at least as they are defined in this paper, start from the assumption that the total electronic energy of a solid can be written in the form

$$E=\sum_{i=1}^{N}\varepsilon_i+\tfrac{1}{2}\sum_{\alpha}\sum_{\beta\,(\neq\alpha)}U(|R_\alpha-R_\beta|) , \tag{14}$$

where the ε_i's are the eigenvalues of some non-self-consistent Schrödinger-like equation,

$$\hat{H}\psi_i(r)=[-\tfrac{1}{2}\nabla^2+V(r)]\psi_i(r)=\varepsilon_i\psi_i(r) , \tag{15}$$

and $U(|R_\alpha-R_\beta|)$ is a short-range pairwise repulsion between the atoms at R_α and R_β.

The origin of Eq. (14) is not at all clear. It looks rather like Eqs. (13c) and (13d), but the double-counting (and nuclear-nuclear repulsion) terms are now assumed to be pairwise and short ranged (which is certainly not the case if charge transfer leads to long-range interatomic Coulomb forces) and the ε_i's are now the solutions of a non-self-consistent Schrödinger equation rather than a self-consistent one. It seems, therefore, that Eq. (14) ignores self-consistency and assumes that all the important nonpairwise behavior in the interatomic forces comes from the sum of the one-electron eigenvalues. In fact, as we will explain, neither of these conclusions is quite right and the approximations behind Eq. (14) are rather more subtle and sophisticated than they appear.

Since the basis of Eq. (14) is not clear it is hard to know how to choose $U(|R_\alpha-R_\beta|)$. It is usually chosen to fit experiment (phonon frequencies, bulk moduli, etc.) and often assumed to have some simple functional form—perhaps an exponential or a power law.

Equation (15) is solved variationally within a basis of localized atomiclike functions, $\{\phi_\alpha\}$, and so leads to a secular equation

$$|\underline{H}-\varepsilon\underline{S}|=0 , \tag{16}$$

where

$$H_{\alpha\beta}=\langle\phi_\alpha|\hat{H}|\phi_\beta\rangle \tag{17a}$$

and

$$S_{\alpha\beta}=\langle\phi_\alpha|\phi_\beta\rangle . \tag{17b}$$

The most common approach is to treat the matrix elements, which are assumed to extend only to first or second neighbors, as disposable parameters and to fit them to experiment or other calculations. In addition, it is often assumed that the basis functions are orthogonal (that they are Wannier functions,[22] for example) and so the overlap matrix $\underline{S}$ is taken to be the unit matrix.

Using the two-center parametrization of Slater and Koster,[23] there are then at least four functions ($ss\sigma$, $sp\sigma$, $pp\sigma$, and $pp\pi$) of interatomic distance to fit for an sp-bonded solid and at least ten ($ss\sigma$, $sp\sigma$, $sd\sigma$, $pp\sigma$, $pp\pi$, $pd\sigma$, $pd\pi$, $dd\sigma$, $dd\pi$, and $dd\delta$) for a solid with s, p, and d electrons. In certain theories (e.g., canonical d-band theory,[24] and Harrison's tight-binding theory[25]) there are relations between these different parameters and rules expressing how they should vary with interatomic distance which serve to cut down the many fitting parameters to a more reasonable number. An alternative approach, favored by Bullett,[26] is to estimate the one-electron potential $V(r)$, choose a basis set, and then do the integrals necessary to evaluate the matrix elements properly. [In fact, Bullett uses the chemical pseudopotentials method and calculates approximations to the $\underline{D}$ matrix elements ($\underline{D}=\underline{S}^{-1}\underline{H}$) rather than to those of $\underline{H}$ and $\underline{S}$, but this is a matter of detail.] His usual estimate for the one-electron potential is obtained by substituting a superposition of atomic densities into the Kohn-Sham form,

$$V(r)=V_{\text{nucl}}(r)+\int\frac{n(r')}{|r-r'|}d^3r'+\mu_{\text{xc}}[n(r)] .$$

A third approach, the linear muffin-tin orbitals (LMTO) tight-binding scheme of Andersen,[27] is not in any sense semiempirical and has no adjustable parameters or functions to be fitted to experiment; however, it is rather more sophisticated than the simple sorts of tight-binding models considered here and cannot easily be related to the framework which has just been presented.

Once choices have been made for the unknown functions it is a straightforward matter to solve the corresponding TB model even for complicated arrangements of atoms. The Hamiltonian matrix can be written straight down and the secular equation treated using the recursion method or by working in **k** space (if there is translational symmetry) or by standard matrix-diagonalization methods. The dimension of the secular equation (number of basis functions per unit cell) is much smaller than in most of the standard (self-consistent) density-functional methods and there is no need for self-consistent looping. Once the secular equation has been solved, the sum of the eigenvalues is added to the pair-potential repulsive energy to give the total energy for that particular arrangement of the frozen atoms.

Changes in total energy as the atoms are moved around (i.e., interatomic forces) and other essentially local quantities can always be cast in terms of local densities of states[7,17] and these can be calculated as functions of the local arrangement of atoms using the recursion method. This approach avoids the need to diagonalize the Hamiltonian matrix for the whole solid and has been successfully used by Sutton, Finnis, Pettifor, and Ohta.[7] It is clearly satisfactory and sensible that local quantities, such as the force on one particular atom, are thus expanded in terms of the local environment.

C. The relationship between density-functional theory and semiempirical tight-binding models

In this section so far, we have given brief introductions to DFT and SETB. Now we proceed to identify the sorts of questions we must answer before we can relate the two.

One obvious group of questions is about the TB energy-functional equation (14), and the non-self-consistent independent-particle equation (15), from which the eigenvalues ε_i are obtained. What has happened to the self-consistent nature of the Kohn-Sham equation, (5)? Can the double-counting energy,

$$F[n_0(r)]-\int \frac{\delta F}{\delta n}\bigg|_{n_0(r)} n_0(r)d^3r \ ,$$

in Eq. (13c) reasonably be represented by a simple short-range repulsive pair potential? And what is the appropriate one-electron potential to use in the tight-binding Hamiltonian, (15)?

The other main group of questions concerns the reduction of Eq. (15) to the TB secular equation (16). What basis functions should one use? Can one assume that they are orthogonal? What about the complicated three-center terms involved in calculating the Hamiltonian matrix? And why should the tight-binding matrix elements be assumed two center and transferable?

In the rest of this paper, we will concentrate on the first of these two groups of questions. In the process, we will also make some headway in understanding the second group, but a fuller discussion of these will be left for a future paper. It turns out that the physics behind Eqs. (14) and (15) can be made much clearer with the use of a little known stationary principle in DFT. This will be derived and explained in the next section.

II. A STATIONARY PRINCIPLE IN DENSITY-FUNCTIONAL THEORY

SETB calculations are clearly not self-consistent and so the best that can be hoped is that they give a variational estimate of the ground-state energy of the solid. In this section we discuss the usual variational principle of DFT and also introduce a less well-known stationary principle which turns out to be a useful starting point from which to try and understand SETB theory. The stationary principle was first used by Wendel and Martin[28] in the course of a non-self-consistent plane-wave and pseudopotential calculation and has since been reintroduced independently by Harris[15] and the present authors.[16]

The total-energy functional in DFT is minimized at the ground-state density $n_0(r)$. It follows that if $\delta n(r)$ is any small normalization-conserving fluctuation in the density,

$$\int \delta n(r)d^3r=0 \ ,$$

then

$$E[n_0+\delta n]=E[n_0]+\tfrac{1}{2}\int\int \frac{\delta^2 E}{\delta n(r)\delta n(r')}\bigg|_{n_0} \delta n(r)\delta n(r')d^3r\, d^3r'+(\text{terms of higher order in } \delta n) \ . \qquad (18)$$

The second-order term, which will be written schematically as

$$\tfrac{1}{2}\int\int \frac{\delta^2 E}{\delta n^2}\bigg|_{n_0} \delta n\, \delta n$$

from now on, is always greater than or equal to zero.

This is the usual variational principle in DFT. It says that if $E[n(r)]$ is evaluated at a guessed ground-state density then the energy obtained will be greater than the ground-state energy by an amount which is second order in the small (we hope) difference between the guessed and exact ground-state densities. The estimated ground-state energy is thus better than the original guess at the ground-state density.

If we look back at Eq. (7), we see that the value of the energy functional at a given density (called n_{out} for reasons which will soon become clear) can be found from the following equations:

$$E[n_{\text{out}}]=T_s[n_{\text{out}}]+F[n_{\text{out}}] \tag{19a}$$

$$=\sum_{i=1}^{N}\int \psi_i^*(r)(-\tfrac{1}{2}\nabla^2)\psi_i(r)d^3r+F[n_{\text{out}}]\ , \tag{19b}$$

$$=\sum_{i=1}^{N}\varepsilon_i-\int V_{\text{in}}(r)n_{\text{out}}(r)d^3r+F[n_{\text{out}}]\ , \tag{19c}$$

where $V_{\text{in}}(r)$ is the potential which makes n_{out} the ground-state density of a noninteracting electron gas and $\psi_i(r)$ is one of the single-particle eigenfunctions of the corresponding independent-electron Hamiltonian.

As has already been mentioned, the usual procedure is to start with an input potential, $V_{\text{in}}(r)$, rather than with a given density. Solving the independent-electron problem then produces the density, $n_{\text{out}}(r)$, at which the functional can be evaluated. If n_{out} is to be close to the ground-state density, then the input potential must have been close to the self-consistent potential and one way of ensuring this is to choose

$$V_{\text{in}}(r)=\left.\frac{\delta F}{\delta n}\right|_{n_{\text{in}}}$$

$$=V_{\text{nucl}}(r)+V_H[n_{\text{in}}(r)]+\mu_{\text{xc}}[n_{\text{in}}(r)]\ , \tag{20}$$

where $n_{\text{in}}(r)$ is a guessed ground-state density.

So the procedure for evaluating $E[n_{\text{out}}(r)]$ is as follows.

(a) Make a guess, $n_{\text{in}}(r)$, at the ground-state density.

(b) Construct the input potential $V_{\text{in}}(r)$ according to Eq. (20).

(c) Solve the separable independent-electron problem to find the density, $n_{\text{out}}(r)$, at which the functional is to be evaluated.

(d) $E[n_{\text{out}}]$ is then given by Eq. (19). With the particular choice for $V_{\text{in}}(r)$ given by Eq. (20) this becomes

$$E[n_{\text{out}}]=\sum_{i=1}^{N}\varepsilon_i-\int\left.\frac{\delta F}{\delta n}\right|_{n_{\text{in}}}n_{\text{out}}+F[n_{\text{out}}] \tag{21a}$$

$$=\sum_{i=1}^{N}\varepsilon_i-\int\int\frac{n_{\text{in}}n_{\text{out}}}{|r-r'|}-\int\mu_{\text{xc}}[n_{\text{in}}]n_{\text{out}}$$

$$+\tfrac{1}{2}\int\int\frac{n_{\text{out}}n_{\text{out}}}{|r-r'|}+E_{\text{xc}}[n_{\text{out}}]\ . \tag{21b}$$

Equations (21) give a strictly variational expression for $E[n_{\text{out}}]$ and have often been used in large self-consistent calculations. However, they involve both n_{in} and n_{out} and are rather more complicated than is convenient for our purposes. It is possible to believe that $V_{\text{in}}(r)$ as given by Eq. (20) is an appropriate potential for use in the tight-binding Hamiltonian [Eq. (15)]; but the output density, $n_{\text{out}}(r)$, does not seem to appear anywhere in the SETB formalism and so there is no immediately obvious relation between Eq. (21b) and the tight-binding estimate of the total energy as given by Eq. (14).

This finishes the discussion of the usual variational principle in DFT. What is not so widely realized is that one can define an alternative energy functional, $\mathcal{E}[n(r)]$, which coincides with $E[n(r)]$ at the ground-state density and which is also stationary about that density. Unlike $E[n(r)]$, however, $\mathcal{E}[n(r)]$ is not strictly variational and so $\mathcal{E}[n_0+\delta n]$ may be either greater than or less than $\mathcal{E}[n_0]$ $(=E[n_0]=E_0)$. For non-self-consistent calculations, the stationary principle (which ensures that the approximate energies are better than the approximate densities from which they were obtained) is the important thing and it does not matter in the least whether the functional is minimized at the stationary point or not. So, unsettling though it may seem, there is no reason to object to the fact that $\mathcal{E}[n]$ is not strictly variational and no reason to expect that energy estimates obtained using $\mathcal{E}[n]$ should be worse than those obtained using $E[n]$.

Now that the usual DFT variational principle has been explained, the stationary functional is very easy to obtain. We will give two derivations which point out different aspects of the physics involved.

The first deviation starts from Eq. (21a). We write

$$\Delta n(r)=n_{\text{out}}(r)-n_{\text{in}}(r)\ ,$$

and then expand $F[n_{\text{out}}]$ about n_{in} to get

$$E[n_{\text{out}}]=\sum_{i=1}^{N}\varepsilon_i+F[n_{\text{in}}]-\int\left.\frac{\delta F}{\delta n}\right|_{n_{\text{in}}}n_{\text{in}}$$

$$+\tfrac{1}{2}\int\int\left.\frac{\delta^2F}{\delta n^2}\right|_{n_{\text{in}}}\Delta n\,\Delta n+\cdots\ . \tag{22}$$

If we now define a new functional, $\mathcal{E}[n_{\text{in}}]$, by

$$\mathcal{E}[n_{\text{in}}]=\sum_{i=1}^{N}\varepsilon_i+F[n_{\text{in}}]-\int\left.\frac{\delta F}{\delta n}\right|_{n_{\text{in}}}n_{\text{in}} \tag{23a}$$

$$=\sum_{i=1}^{N}\varepsilon_i-E_H[n_{\text{in}}]+E_{\text{xc}}[n_{\text{in}}]-\int\mu_{\text{xc}}[n_{\text{in}}]n_{\text{in}}\ , \tag{23b}$$

then Eq. (22) becomes

$$\mathcal{E}[n_{\text{in}}]=E[n_{\text{out}}]-\tfrac{1}{2}\int\int\left.\frac{\delta^2F}{\delta n^2}\right|_{n_{\text{in}}}\Delta n\,\Delta n$$

$$+(\text{higher-order terms})\ . \tag{24}$$

$\mathcal{E}[n_{\text{in}}]$ is the functional we want: it differs from $E[n_{\text{out}}]$ by terms which are second order in Δn and is clearly stationary about the ground-state density; and when evaluated at that density, it is the same as the usual Hohenberg-Kohn[3] functional, $E[n_0]$, and so is equal to the ground-state energy. In addition, Eq. (23b) shows that $\mathcal{E}[n_{\text{in}}]$ does not depend explicitly on n_{out}; the eigenvalues are the solutions of a non-self-consistent Schrödinger equation with a potential constructed using n_{in} alone,

$$\{-\tfrac{1}{2}\nabla^2+V_{\text{nucl}}(r)+V_H[n_{\text{in}}(r)]$$

$$+\mu_{\text{xc}}[n_{\text{in}}(r)]\}\psi_i(r)=\varepsilon_i\psi_i(r)\ ;$$

and the "double-counting" terms also depend solely on n_{in},

$$F[n_{\text{in}}]-\int \left.\frac{\delta F}{\delta n}\right|_{n_{\text{in}}} n_{\text{in}}$$

$$=-E_H[n_{\text{in}}]+E_{\text{xc}}[n_{\text{in}}]-\int \mu_{\text{xc}}[n_{\text{in}}]n_{\text{in}} \ .$$

$\mathcal{E}[n_{\text{in}}]$ can therefore be evaluated without needing to know the one-electron eigenfunctions or the output density and is obviously a sensible starting point from which to attempt a derivation of something very similar to the SETB equations, (14) and (15).

An alternative derivation of the stationary functional starts from the Hohenberg-Kohn variational estimate evaluated at n_{in},

$$E[n_{\text{in}}]=T_s[n_{\text{in}}]+F[n_{\text{in}}] \ .$$

The difficulty in evaluating this functional stems from the fact that we do not know the potential needed to make n_{in} the ground-state density of the noninteracting system and so cannot evaluate $T_s[n_{\text{in}}]$. However, if we expand $T_s[n_{\text{in}}]$ about n_{out} and use Eq. (8) we get

$$T_s[n_{\text{in}}]=T_s[n_{\text{out}}]+\int V_{\text{in}}\,\Delta n$$
$$+\tfrac{1}{2}\int\int \left.\frac{\delta^2 T_s}{\delta n^2}\right|_{n_{\text{out}}} \Delta n\,\Delta n+\cdots \ .$$

Using Eq. (7b) for $T_s[n_{\text{out}}]$ in terms of the one-electron eigenvalues and the input potential then gives

$$T_s[n_{\text{in}}]=\sum_{i=1}^{N}\varepsilon_i-\int V_{\text{in}}n_{\text{in}}$$
$$+\tfrac{1}{2}\int\int \left.\frac{\delta^2 T_s}{\delta n^2}\right|_{n_{\text{out}}} \Delta n\,\Delta n+\cdots$$

and hence

$$E[n_{\text{in}}]=\sum_{i=1}^{N}\varepsilon_i+F[n_{\text{in}}]-\int V_{\text{in}}n_{\text{in}}+\tfrac{1}{2}\int\int \left.\frac{\delta^2 T_s}{\delta n^2}\right|_{n_{\text{out}}} \Delta n\,\Delta n+\cdots$$
$$=\sum_{i=1}^{N}\varepsilon_i+F[n_{\text{in}}]-\int \left.\frac{\delta F}{\delta n}\right|_{n_{\text{in}}} n_{\text{in}}+\tfrac{1}{2}\int\int \left.\frac{\delta^2 T_s}{\delta n^2}\right|_{n_{\text{out}}} \Delta n\,\Delta n+\cdots$$
$$=\mathcal{E}[n_{\text{in}}]+\tfrac{1}{2}\int\int \left.\frac{\delta^2 T_s}{\delta n^2}\right|_{n_{\text{out}}} \Delta n\,\Delta n+\cdots \ .$$

So we now have two alternative expressions for $\mathcal{E}[n_{\text{in}}]$ correct to second order in Δn:

$$\mathcal{E}[n_{\text{in}}]=E[n_{\text{out}}]-\tfrac{1}{2}\int\int \left.\frac{\delta^2 F}{\delta n^2}\right|_{n_{\text{in}}} \Delta n\,\Delta n \tag{25a}$$

$$=E[n_{\text{in}}]-\tfrac{1}{2}\int\int \left.\frac{\delta^2 T_s}{\delta n^2}\right|_{n_{\text{out}}} \Delta n\,\Delta n \ . \tag{25b}$$

To derive the first of these we expanded the double-counting part of $E[n_{\text{out}}]$ to second order in $n_{\text{out}}-n_{\text{in}}$, whereas to derive the second we expanded the kinetic energy part of $E[n_{\text{in}}]$ to second order in $n_{\text{out}}-n_{\text{in}}$. The equivalence of Eqs. (25a) and (25b) is not immediately obvious, but follows from the fact that n_{out} satisfies a noninteracting Schrödinger equation with an input potential constructed using n_{in}.

Equations (25a) and (25b) can be combined to give a third expression which shows, perhaps more clearly, the nature of the approximation being made. Let us write

$$\Delta n_{\text{in}}=n_{\text{in}}-n_0 \quad\text{and}\quad \Delta n_{\text{out}}=n_{\text{out}}-n_0 \ .$$

Then $\Delta n=\Delta n_{\text{out}}-\Delta n_{\text{in}}$. Adding (25a) to (25b) and working to second order in the small quantities Δn_{out} and Δn_{in} gives

$$\mathcal{E}[n_{\text{in}}]=E_0+\tfrac{1}{2}\int\int \left.\frac{\delta^2 E}{\delta n^2}\right|_{n_0} \Delta n_{\text{in}}\,\Delta n_{\text{out}}+\cdots \ . \tag{26}$$

The functional derivative here is the derivative of the ordinary Hohenberg-Kohn functional evaluated at the ground-state density and so Eq. (26) invites comparison with the corresponding expressions for the Hohenberg-Kohn functional at n_{in} and n_{out}:

$$E[n_{\text{in}}]=E_0+\tfrac{1}{2}\int\int \left.\frac{\delta^2 E}{\delta n^2}\right|_{n_0} \Delta n_{\text{in}}\,\Delta n_{\text{in}}+\cdots \ ,$$

$$E[n_{\text{out}}]=E_0+\tfrac{1}{2}\int\int \left.\frac{\delta^2 E}{\delta n^2}\right|_{n_0} \Delta n_{\text{out}}\,\Delta n_{\text{out}}+\cdots \ .$$

Both $E[n_{\text{in}}]$ and $E[n_{\text{out}}]$ are greater than or equal to E_0 because the second functional derivative of the Hohenberg-Kohn functional is non-negative. However, $\mathcal{E}[n_{\text{in}}]$ differs from E_0 by a term depending on both Δn_{in} and Δn_{out} and so this is not enough to ensure that $\mathcal{E}[n_{\text{in}}]$ is also greater than E_0. This shows clearly why $\mathcal{E}[n]$is a stationary functional but not a variational one. In fact, since it is common to see "overshoot" during the first few iterations of a self-consistent calculation (so that the errors in n_{in} are more than compensated in n_{out}), we expect that $\mathcal{E}[n_{\text{in}}]$ may often be less than the ground-state energy. This will be demonstrated in an example later on.

We finish this section with one final observation. In the second derivation of the stationary functional, it was assumed that V_{in} was constructed by substituting n_{in} in to the usual Kohn-Sham form, Eq. (20). However, it is clear that this was not a necessary assumption: all that mattered was that n_{out} was the density obtained by solving the independent-electron problem with an input potential equal to V_{in}. It is only when V_{in} is related to n_{in} according to Eq. (20) that the stationary approximation takes the form given in Eqs. (23a) and (23b), but if we define $\mathcal{E}[n_{\text{in}},V_{\text{in}}]$ according to

$$\mathcal{E}[n_{\text{in}},V_{\text{in}}]=\sum_{i=1}^{N}\varepsilon_i+F[n_{\text{in}}]-\int V_{\text{in}}n_{\text{in}} ,$$

then it is always stationary. This extra freedom may be helpful in some applications (whenever it is necessary to make some sort of shape approximation for the potential, for example), but will not be used here. The equivalent of Eq. (26) for this more general functional is

$$\mathcal{E}[n_{\text{in}},V_{\text{in}}]=E_0+\tfrac{1}{2}\int\int\frac{\delta^2E}{\delta n^2}\bigg|_{n_0}\Delta n_{\text{out}}\,\Delta n_{\text{in}} -\tfrac{1}{2}\int(V[n_{\text{in}}]-V_{\text{in}})\Delta n ,$$

where $V[n_{\text{in}}]$ is the usual Kohn-Sham potential as given by Eq. (20).

III. THE TIGHT-BINDING ENERGY EXPRESSION

In this section, we show how the use of the stationary principle with an appropriate guessed input density leads directly to an energy expression of the sort assumed in SETB theory.

The value of the stationary functional at a density $n(r)$ (the subscripts "in" and "out" are no longer needed as the stationary functional only involves n_{in}) is given by

$$\mathcal{E}[n]=\sum_{i=1}^{N}\varepsilon_i+F[n(r)]-\int\frac{\delta F}{\delta n}\bigg|_{n(r)}n(r)d^3r \tag{27a}$$

$$=\sum_{i=1}^{N}\varepsilon_i-E_H[n(r)]-\int\mu_{\text{xc}}[n(r)]n(r)d^3r +E_{\text{xc}}[n(r)] , \tag{27b}$$

where the eigenvalues, ε_i, are the solutions of a non-self-consistent independent-electron problem of the form

$$[-\tfrac{1}{2}\nabla^2+V(r)]\psi_i(r)=\varepsilon_i\psi_i(r) , \tag{28}$$

and the input potential, $V(r)$, is constructed from $n(r)$ according to

$$V(r)=\frac{\delta F}{\delta n}\bigg|_{n(r)}=V_{\text{nucl}}(r)+V_H[n(r)]+\mu_{\text{xc}}[n(r)] . \tag{29}$$

Before doing anything else, let us discuss the frozen-core approximation,[29] which turns out to be particularly simple when expressed in terms of the stationary functional and enables us, if we choose, to get away without considering the core electrons explicitly. The assumption is that the core-electronlike solutions, ψ_{ic}, of Eq. (28) are very close to linear combinations of nonoverlapping unperturbed atomic core functions, $\phi_{\alpha c}$. Since the core functions are well localized around the nuclei, and since the core electrons in the deep potential well near one nucleus "feel" the effects of the other atoms as comparatively weak perturbations, this seems a sensible enough approximation. At any rate, the frozen-core approximation underlies all pseudopotential calculations and its accuracy has been well demonstrated in practice.

Suppose the differences between the core states, ψ_{ic}, and appropriate Bloch sums of nonoverlapping atomic core orbitals, $\phi_{\alpha c}$, are of order λ, where λ is small. Then the ordinary variational principle of one-electron quantum theory tells us that

$$\sum_{i,c}\varepsilon_{ic}=\sum_{i,c}\langle\psi_{ic}|[-\tfrac{1}{2}\nabla^2+V(r)]|\psi_{ic}\rangle$$
$$=\sum_{\alpha,c}\langle\phi_{\alpha c}|[-\tfrac{1}{2}\nabla^2+V(r)]\phi_{\alpha c}\rangle+O(\lambda^2)$$
$$=T_c+\int n_cV+O(\lambda^2) ,$$

where T_c is the total kinetic energy of a set of unperturbed atomic cores within DFT and n_c is the corresponding sum of atomic core densities. Equation (27b) then becomes (ignoring terms second order or higher in λ)

$$\mathcal{E}[n]=\sum_{i,v}\varepsilon_{iv}+T_c+\int n_cV-E_H[n] -\int\mu_{\text{xc}}[n]n+E_{\text{xc}}[n] .$$

Writing $n=n_c+n_v$ (since n_c is known, this equation defines n_v given n or vice versa), adding the nuclear-nuclear terms, and using the fact that n_c is a sum of nonoverlapping spherically symmetric atomic core densities then gives

$$\mathcal{E}[n]=\sum_{i,v}\varepsilon_{iv}-E_H[n_v]-\int\mu_{\text{xc}}[n]n_v+E_{\text{xc}}[n] +\tfrac{1}{2}\sum_{\alpha}\sum_{\beta(\neq\alpha)}\frac{Z_{v\alpha}Z_{v\beta}}{|R_\alpha-R_\beta|}+C , \tag{30}$$

where C is a structure-independent constant (except for the exchange and correlation terms, C is just the sum of the free atomic core energies) and will be ignored from now on. $Z_{v\alpha}$ is the number of valence electrons on the atom at R_α and the rest of the nuclear-nuclear repulsion has been screened out by the frozen-core electrons. The evaluation of the left-hand side of Eq. (30) requires only the valence eigenvalues and these are the solutions of an independent-particle equation of the form

$$\{-\tfrac{1}{2}\nabla^2+V_{\text{nucl}}(r)+V_H[n_c(r)]+V_H[n_v(r)] +\mu_{\text{xc}}[n_c(r)+n_v(r)]\}\psi_{iv}(r)=\varepsilon_{iv}\psi_{iv}(r) . \tag{31}$$

Equations (30) and (31) together comprise the frozen-core approximation.

We can now replace the one-electron potential due to the frozen cores by a sum of norm-conserving ionic pseudopotentials.[30] The independent-electron problem is then

$$\left[-\tfrac{1}{2}\nabla^2+\sum_{\alpha}V_{\text{ps},\alpha}+V_H[n_v]+\mu_{\text{xc}}[n_v]\right]\psi_{\text{ps},iv}=\varepsilon_{iv}\psi_{\text{ps},iv} , \tag{32}$$

and the corresponding energy expression is

$$\mathcal{E}[n]=\sum_{i,v}\varepsilon_{iv}-E_H[n_v]-\int\mu_{\text{xc}}[n_v]n_v+E_{\text{xc}}[n_v]+\tfrac{1}{2}\sum_{\alpha}\sum_{\beta\,(\neq\alpha)}\frac{Z_{v\alpha}Z_{v\beta}}{|R_\alpha-R_\beta|}\ . \tag{33}$$

The approximations involved in this step (the replacement of the full frozen-core potential by a pseudopotential and an implicit "linearization" of the exchange and correlation functional) are not straightforward but are well established and have been discussed by Louie, Froyen, and Cohen[31] among others.

So, the frozen-core approximation tells us that as long as we replace the nuclear Coulomb potential by a sum of frozen-core potentials (or pseudopotentials) and remember that the frozen-core electrons screen out most of the nuclear-nuclear Coulomb interaction, we can ignore the core electrons altogether (to second order in λ) and only need to solve the Schrödinger-like Kohn-Sham equations for the valence electrons. This simplification is particularly useful when the frozen-core potentials are replaced by norm-conserving pseudopotentials, since the valence eigenenergies are then the lowest eigenvalues of the independent-electron problem and can be calculated without prior knowledge of the core functions. [In fact, explicit calculations[29] show that the assumption that the core densities are frozen is not a very good one. The frozen-core approximation only works because the stationary principle acts effectively (i.e., λ^2 is small even though λ is not so small) and forces the frozen-core energies to be better than the frozen-core densities from which they were generated.]

In the rest of this section, we will assume that we are working within a pseudopotential framework and will not consider the core electrons. To regain the all-electron results from the ones given here, just replace n_v by n, Z_v by Z, and V_{ps} by V_{nucl} throughout.

It is now necessary to choose $n_v(r)$, the input valence density from which the input potential in Eq. (32) will be constructed and at which the stationary functional will be evaluated. As explained earlier, we seek an input valence density close to the true ground-state valence density. There are many different possible guesses for $n_v(r)$ and the choice must be made on physical grounds and will depend on the solid studied. For a weak pseudopotential metal, such as aluminum, it might be best to generate n_v by treating the effect of the pseudopotential on a uniform electron gas in the linear screening approximation. For nickel, it would probably be better to use a superposition of atomic densities. For carbon, a superposition of typical molecular bond densities might be sensible. Note that because $\mathcal{E}[n]$ is stationary about the ground-state density, one can leave a few variable parameters in the form of the input density and adjust them during the calculation to find the stationary point and, hence, optimize the density.

Overall, it should be possible to come up with plausible approximate densities for most cases of interest. Some knowledge of the solid is needed but this can be obtained from molecular calculations or full self-consistent DFT calculations. Whether the stationary principle will then prove effective enough to produce reliable ground-state energies from these fairly crude guessed densities remains to be seen.

For the rest of this paper, we will work with input densities which are simple superpositions of spherical atomiclike (not necessarily atomic) densities. These seem to work fairly well on the whole, but it is important to stress that they are just the first and simplest approximation and that it is possible to do considerably better in many cases. It is clear that attempting to superpose atomic densities might not be a sensible idea in highly ionic solids, for example (although see the work of Polatoglou and Methfessel[32] on NaCl).

At any rate, let us suppose for now that the guessed input valence density is in the form

$$n_v(r)=\sum_\alpha n_{v\alpha}(r)\ ,$$

where $n_{v\alpha}(r)$ is a spherical density on the atom at R_α and has probably been obtained from some sort of atomic calculation. Given this choice for the input density, the first thing to notice is that the "double-counting" contribution to the stationary functional is very nearly just a simple pairwise function of the interatomic distances and is straightforward to calculate.

The "electrostatic" part of the double-counting energy plus the nuclear-nuclear repulsion is strictly pairwise and is short ranged if the input density is chosen so that the atoms are neutral. It is given by

$$\begin{aligned}D_c[n_v]&=\tfrac{1}{2}\sum_\alpha\sum_{\beta\,(\neq\alpha)}\frac{Z_{v\alpha}Z_{v\beta}}{|R_\alpha-R_\beta|}-\tfrac{1}{2}\int\int\frac{n_v(r)n_v(r')}{|r-r'|}d^3r\,d^3r'\\&=A+\tfrac{1}{2}\sum_\alpha\sum_{\beta\,(\neq\alpha)}\left[\frac{Z_{v\alpha}Z_{v\beta}}{|R_\alpha-R_\beta|}-\int\int\frac{n_{v\alpha}(r)n_{v\beta}(r')}{|r-r'|}d^3r\,d^3r'\right]\ ,\end{aligned}\tag{34}$$

where the constant A is just minus the intra-atomic Hartree energy. If charge transfer is allowed, there is a long-ranged contribution in the form of a standard Madelung sum.

The exchange and correlation (xc) contributions are not quite so simple. Within the local-density approximation they can be written

$$D_{\mathrm{xc}}[n_v]=\int[\varepsilon_{\mathrm{xc}}(n_v(r))-\mu_{\mathrm{xc}}(n_v(r))]n_v(r)d^3r \ . \tag{35}$$

This is not exactly reducible to a pair potential because $\varepsilon_{\mathrm{xc}}$ and μ_{xc} are nonlinear functions of their arguments. However, if there are only small regions in the solid where the overlap of densities from three or more atoms is appreciable, D_{xc} can be approximated by a cluster expansion

$$D_{\mathrm{xc}}\left[\sum_\alpha n_{v\alpha}\right]=\sum_\alpha D_{\mathrm{xc}}[n_{v\alpha}]+\tfrac{1}{2}\sum_\alpha\sum_{\beta\,(\neq\alpha)}(D_{\mathrm{xc}}[n_{v\alpha}+n_{v\beta}]-D_{\mathrm{xc}}[n_{v\alpha}]-D_{\mathrm{xc}}[n_{v\beta}]) \\ +(\text{three- and higher-body interactions}) \ . \tag{36}$$

Estimates suggest that the errors involved in truncating this expansion after the second (two-body) term are usually very small and so the xc double-counting contributions are also well represented by a simple pair potential.

So we reach the happy conclusion that it will be a good approximation to replace the double-counting and nuclear-nuclear repulsion energy as a whole by a simple sum of one- and two-body interatomic potentials: the double-counting energy is equal to

$$\sum_\alpha C_\alpha+\tfrac{1}{2}\sum_\alpha\sum_{\beta\,(\neq\alpha)}U_{\alpha\beta}(|R_\alpha-R_\beta|) \ , \tag{37}$$

where C_α is the double-counting energy for an isolated monomer and $U_{\alpha\beta}$ is a pair potential which can be found from a dimer calculation. As long as the input density is a superposition of spherical atomiclike densities, both C_α and $U_{\alpha\beta}$ are easily calculable (it is necessary to evaluate a few three-dimensional integrals to find them) and are completely transferable. They can both be calculated once and for all (from monomer and dimer calculations, respectively) and stored for future use.

The other contribution to the stationary total energy, Eq. (33), is the sum of the one-electron eigenvalues obtained from the independent-particle Schrödinger equation (32). If the input density is a superposition of spherical atomiclike densities, then the one-electron potential in this equation is of the form

$$V(r)=\sum_\alpha V_\alpha(r)+U(r) \ , \tag{38}$$

where V_α is the full screened [i.e., including the Hartree and xc contributions from the valence density $n_{v\alpha}(r)$] atomic pseudopotential on the atom at R_α,

$$V_\alpha(r)=V_{\mathrm{ps},\alpha}(r)+V_H[n_{v\alpha}(r)]+\mu_{\mathrm{xc}}[n_{v\alpha}(r)] \ , \tag{39}$$

and $U(r)$ is a small extra potential due to the nonlinearity of the xc functional,

$$U(r)=\mu_{\mathrm{xc}}\left[\sum_\alpha n_{v\alpha}(r)\right]-\sum_\alpha\mu_{\mathrm{xc}}[n_{v\alpha}(r)] \ . \tag{40}$$

So as well as leading to a double-counting energy which is almost pairwise, superposing atomiclike densities produces an input one-electron potential of a particularly simple form. Other choices for the input density do not lead to quite such simple results, but their analysis and implementation is not much more difficult.

Looking back, it can now be seen that all the questions about the forms of the SETB energy expression, Eq. (14), and independent-electron equation, (15), have been answered. It has been shown why it is possible in many cases to approximate the double-counting part of the stationary functional as a simple sum of transferable one- and two-body potentials [Eqs. (34)–(37)]; and an explicit construction [Eqs. (38)–(40)] has been given for the one-electron potential to be used in the independent-electron equation. We have not answered all the questions about the reduction of Eq. (15) to a secular equation, but at least the prescription for generating the input potential implies that the value of $V(r)$ at any point depends only on the local environment. This is a prerequisite if we are to end up with Hamiltonian matrix elements depending on the positions of just a few atoms, although it is not enough to ensure that they are simple pairwise functions of interatomic distance and does not tell us how to choose an appropriate basis set. We return to consider these questions in the concluding section of this paper.

All the results in this section rely, of course, on the assumption that we will be able to guess the input density well enough to make the energy estimates obtained using the stationary functional accurate enough to be useful. If we are to attempt defect calculations, we will have to be able to do this for a wide variety of different arrangements of the atoms. In the next section we make some preliminary tests of this assumption.

IV. TESTS OF THE STATIONARY EXPRESSION FOR THE TOTAL ENERGY

In this section, we test the stationary energy expression with superimposed atomiclike input densities in a few simple cases. The simplest interesting cases are dimers and we will look at a selection of these involving hydrogen, helium, and germanium atoms. In fact, dimers make better test cases than might at first be imagined: although each atom has only one neighbor, the bond to that neighbor is often strong, the interatomic distance often small, and the difference between superposed atomic densities and the true density is often comparable to that in a solid. Harris[15] has also looked at some dimers using the stationary functional and it is interesting to compare his results with ours.

Given an input density which is a superposition of spherical atomiclike densities for the two atoms, it is easy enough to work out the double-counting terms and the effective one-electron Hamiltonian (we use the parametrization of Perdew and Zunger[21] for the exchange and correlation functional in the local-density approximation) for the dimer. We then solve the independent-electron problem variationally in a basis of localized atomiclike orbitals. Given that our overall aim is to calculate transferable sets of tight-binding parameters ($ss\sigma$, $sp\sigma$, etc.) for later use in calculations of the properties of solids, it is clearly necessary that we use such a basis and that the physics be not too sensitive to the exact basis functions chosen. Hence, this aspect of the calculation also provides a useful test (although not of the functional itself). Note that there is no need to relate the basis functions to the input density used in the stationary functional, and so although we may choose to use exact atomic densities in the functional this does not mean that the basis functions must also be atomic. Plane waves would be equally acceptable and may well be better in many cases. Note also that the choice of a finite basis set at this stage is a strictly variational (and not just stationary) approximation and so the better the basis set the lower the sum of the occupied one-electron eigenvalues.

Once the basis functions have been chosen, the evaluation of the Hamiltonian and overlap matrices for the dimer is a simple matter of doing a few three-dimensional integrals (reducible to two-dimensional integrals by symmetry). Most of these are in the form of convolutions and may be evaluated by using the convolution theorem to rewrite them in terms of one-dimensional spherical Hankel transforms and Gaunt coefficients. [The ℓth spherical Hankel transform, $\mathscr{f}_\ell(k)$, of a one-dimensional function $f(r)$ is given by

$$\mathscr{f}_\ell(k)=\int_0^\infty f(r)j_\ell(kr)r^2\,dr ,$$

where j_ℓ is a spherical Bessel function. A Gaunt coefficient is an integral of a product of three spherical harmonics.] In our use of this technique, the Hankel transforms were evaluated numerically with an accuracy of around 10^{-5} hartree and the Gaunt coefficients were obtained using a standard library subroutine. The matrix elements of $U(r)$ (the part of the potential due to the nonlinearity of the exchange-correlation functional) and the corresponding terms in the double-counting energy cannot be evaluated using the convolution trick and so a two-dimensional Gauss-Chebyshev quadrature method as developed by Nex[33] was employed.

Once the Hamiltonian and overlap matrix elements have been evaluated, the secular equation can be written straight down and, of course, is block diagonal in the azimuthal quantum number, M. It is easily solved using standard matrix eigenvalue routines or by hand in many cases. The sum of occupied eigenvalues is then added to the nuclear-nuclear Coulomb repulsion and the double-counting "pair-potential" energy to give the value of the stationary functional at the input density. The calculation is repeated at different interatomic spacings (and possibly at different input densities and with different basis sets) to find the dimer energy as a function of interatomic spacing and hence the equilibrium bond length, binding energy, and vibrational frequency.

In comparison with accurate local-density (pseudopotential) calculations, this scheme involves two extra approximations, one of which is stationary and one of which is variational. There is the use of the stationary functional evaluated at an input density which is a superposition of atomiclike densities; and there is the use of a finite basis of localized functions in the solution of the eigenproblem.

A. The hydrogen atom

This is almost the simplest quantum-mechanical system imaginable and the simplest on which we can sensibly test the stationary functional. There are no core electrons and little point in attempting to construct a pseudopotential. The Schrödinger equation is

$$\left[-\tfrac{1}{2}\nabla^2-\frac{1}{r}\right]\psi(r)=E\psi(r) , \tag{41}$$

and this has a ground-state wave function

$$\psi_0(r)=\frac{1}{\sqrt{\pi}}e^{-r} ,$$

with ground-state energy -0.5 hartree.

The density-functional theory equations for the H atom are

$$\left[-\tfrac{1}{2}\nabla^2-\frac{1}{r}+V_H[n_0(r)]+\mu_{\rm xc}[n_0(r)]\right]\psi_0(r)=\varepsilon_0\psi_0(r) ,$$

$$E[n_0(r)]=\varepsilon_0-E_H[n_0(r)]+E_{\rm xc}[n_0(r)] - \int\mu_{\rm xc}[n_0(r)]n_0(r)d^3r ,$$

where

$$n_0(r)=|\psi_0(r)|^2 .$$

In exact density-functional theory these would give exactly the same density and energy as Eq. (41) and so, in this simple case, the exact exchange and correlation functionals must be given by

$$E_{\rm xc}[n_0(r)]=-E_H[n_0(r)] ,$$

$$\mu_{\rm xc}[n_0(r)]=-V_H[n_0(r)]+\text{const} .$$

Within the LDA these equations are far from satisfied and so it is perhaps surprising that self-consistent LDA calculations give sensible results for the hydrogen atom. They do, however, as was demonstrated by Gunnarsson, Lundqvist, and Wilkins,[34] who found a total energy of -0.492 hartree and an electron density very little different from the exact one.

Before going on to discuss the results obtained using the approximate functional, it is necessary to point out one complication which arises in atomic calculations. Hund's first rule says that the electrons in an atom will occupy the one-electron energy levels in such a way as to maximize the total spin S. This means that most atoms show some spin polarization and, because the Pauli principle only keeps apart electrons with the same spin, this affects the total exchange and correlation energy. As a result, it is necessary to go beyond the LDA and use the spin-dependent version, the local-spin-density (LSD) approximation, in most atomic calculations. This is rarely necessary in (nonmagnetic) solids, however, and so LDA calculations will suffice in most cases of interest to us. Here we content ourselves with including spin-dependent effects in a very approximate way which we now describe.

Within the Hartree-Fock (i.e., exchange-only) approximation, the exchange energy density of a uniform electron gas with spin-up and spin-down densities n_+ and n_- is given by

$$-\frac{3}{4}\left[\frac{3}{\pi}\right]^{1/3} 2^{1/3}(n_+^{4/3}+n_-^{4/3}) \; . \tag{42}$$

In the spin-compensated case, when $n_+=n_-=n/2$, the corresponding local density approximation would be

$$\varepsilon_{\mathrm{xc}}[n]=-\frac{3}{4}\left[\frac{3}{\pi}\right]^{1/3} n^{1/3} \; . \tag{43}$$

Equation (42) says that the exchange energy in a perfectly spin-polarized electron gas ($n_+=n$, $n_-=0$) is $2^{1/3}$ times that in a spin-compensated gas of the same density. So an approximate way of taking account of the spin polarization in the hydrogen atom is to multiply the LDA xc energy by $2^{1/3}$. This is a crude approximation, of course, but since we are not really very interested in spin-polarized situations it will do for now.

Using this approximation, an input density equal to the exact hydrogenic ground-state density, and a single exact hydrogenic 1*s* wave function as the basis set, gave a binding energy of 0.507 hartree. That this is slightly different from the result of Gunnarsson *et al.* could be due to the approximate handling of the spin-polarization energy or to the fact that the input density, although exact for a real hydrogen atom, is not exact for a self-consistent LDA hydrogen atom. For comparison, a similar calculation without the approximate correction for spin polarization gave a binding energy of only 0.378 hartree.

B. The hydrogen molecule

As two hydrogen atoms are brought together, the spin polarization of each gradually decreases until, well before the equilibrium spacing is reached, the bonding *s* orbital contains one up and one down electron and the spin density is zero. The properties of the dimer near the equilibrium spacing should therefore be calculable within the LDA and there is no need to worry about spin polarization.

In order to check the reliability of the stationary energy functional several different input densities were tried.

Density 1: A superposition of exact atomic densities.

Density 2: A superposition of spherical densities of the form

$$n_{\mathrm{sph}}(r)=\frac{0.61}{\pi}e^{-2r}+\left[\frac{2.2}{\pi}\right]^{3/2}(1-0.61)e^{-2.2r^2} \; .$$

Density 3: A superposition of spherical densities of the form

$$n_{\mathrm{sph}}(r)=\frac{0.4}{\pi}e^{-2r}+\left[\frac{2.2}{\pi}\right]^{3/2}(1-0.4)e^{-2.2r^2} \; .$$

Figure 1 shows a graph of input densities 1 and 2 and of the accurate dimer density[35] along the internuclear axis at the equilibrium spacing of $1.4a_0$ ($1a_0=1$ Bohr radius). It is worth noting that the superposition of atomic densities is almost 40% smaller than the true density at the center of the bond and so we are putting the stationary energy functional to a severe test with these calculations.

Numerical experiments showed that an adequate basis set consists of just two nodeless *s* functions on each hydrogen atom. One of these is an atomic *s* function (Y_{lm} is a spherical harmonic),

$$\phi_1(r)=2e^{-r}Y_{00}(\hat{\mathbf{r}}) \; ,$$

and so ensures that (bar the spin-polarization problem) the calculation gives good answers as the two atoms are pulled far apart; and the other is of the form

$$\phi_2(r)=2(1.3)^{3/2}e^{-1.3r}Y_{00}(\hat{\mathbf{r}}) \; ,$$

which is the variationally best simple exponential 1*s*-like orbital at the equilibrium interatomic spacing. It was found that adding 2*s* and 2*p* hydrogenic orbitals to this basis set made little difference to the results of the calculations.

The results for the binding energy, interatomic spacing, and vibrational frequency of H_2 are shown in Table I. Included for comparison are the almost exact configuration-interaction results of Kolos and Rooth-

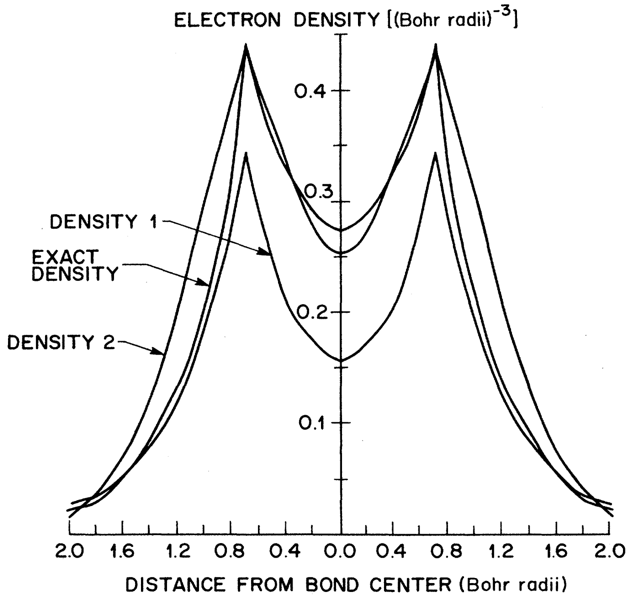

FIG. 1. Comparison of the exact electron density and the different input densities along the molecular axis for a hydrogen dimer at equilibrium spacing (for definitions of densities 1 and 2, see text).

TABLE I. Results of various calculations of the binding energy, E_b, the equilibrium interatomic spacing, R_e, and the vibrational frequency, w_e, for H_2.

	E_b (eV)	R_e/a_0	w_e (cm^{-1})
Kolos and Roothaan[a]	4.75	1.40	4410
Gunnarsson and Johansson[b]	4.79	1.44	
Painter[c]	4.90 (approx.)	1.44	4000
Density 1	5.11	1.38	4220
Density 2	4.89	1.49	3830
Density 3	5.24	1.50	3830

[a] Reference 36.
[b] Reference 37.
[c] Reference 38.

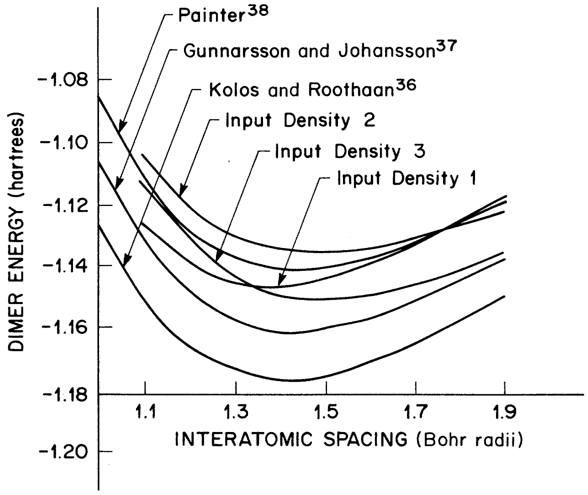

FIG. 2. Total energy as a function of interatomic spacing for the hydrogen dimer.

aan[36] and the accurate DFT calculations of Gunnarsson and Johansson[37] (GJ) and of Painter.[38]

Gunnarsson and Johansson use the form of the LDA due to Gunnarsson and Lundqvist (GL),[39] whereas Painter uses that due to Vosko, Wilk, and Nusair (VWN).[40] As demonstrated by Painter, the calculated binding energy is slightly larger with the VWN xc potential than with the GL one. The xc potential used here, due to Perdew and Zunger,[21] is based, like the VWN one, on the electron-gas Monte Carlo calculations of Ceperley and Alder,[20] and so we might expect that our results will be closer to those of Painter than to those of GJ. Since we did not do a proper spin-polarized calculation for the atoms, our binding energies were calculated as the difference between the total energy of the dimer at the equilibrium spacing and twice Painter's value of the atomic energy.

As Table I shows, the results are very encouraging. The binding energies lie within about 10% of Painter's values and the interatomic spacings and vibrational frequencies are within about 5%. Considering the very large differences between the various input densities and the exact density, the effectiveness of the stationary property of the approximate functional is clearly demonstrated.

Graphs of the total energy against interatomic spacing are shown in Fig. 2 (the results of other workers were copied from diagrams in papers and are, therefore, not very accurate). As one might expect given the stationary but not variational nature of the approximate functional, it is possible to obtain total energies either greater than or less than Painter's accurate value calculated with a very similar xc potential. It is reassuring to see that the errors due to the use of the approximate functional are no bigger than those due to the use of the LDA, although they are considerably more dependent on interatomic spacing and so affect the physical properties more.

C. The helium dimer

Having looked at H_2, which is a strongly bound open-shell system, it seems sensible next to look at a closed-shell system such as He_2. The binding energy of He_2 is only about 1 meV and since the binding is the result of the van der Waals interaction one should not expect to calculate it accurately using a local-density approximation (although, see Harris[15]). However, the strong closed-shell repulsion at smaller interatomic distances should be calculable and it is this at which we will look.

It was decided to use input densities in the form of simple exponentials with decay rates chosen by comparison with the results of self-consistent $X\alpha$ calculations for the helium atom (Herman and Skillman[41]). Two different decay rates were tried:

$$n_1(r)=\frac{2(1.94)^3}{\pi}e^{-2(1.94)r},$$

$$n_2(r)=\frac{2(1.55)^3}{\pi}e^{-2(1.55)r}.$$

Density 1 was chosen because it has about the right decay rate in the region $r=0.1a_0$ to $r=0.5a_0$; and density 2 because it gives more or less the right electron density near $r=0$. In neither case is the fit to the Herman-Skillman density very good, but it turns out that the stationary principle acts very efficiently in this case and even these rather poor input densities are quite satisfactory.

For both input densities, the basis set consisted of just one s orbital per atom and this was chosen so that, when doubly occupied, it gave rise to the corresponding input density. There are no core electrons and so the full potential was used rather than a pseudopotential.

The results were compared with the configuration-interaction calculations of Phillipson,[42] as no DFT calculations extending over a wide enough range of interatomic spacings were found.

The atomic energies differed quite considerably in the three calculations. Phillipson's value is about −2.875 hartrees, density 1 gave −2.73 hartrees, and density 2 gave −2.80 hartrees. This shows that the input densities and basis functions were not very good. However, as shown in Fig. 3, the repulsion between the two atoms is remarkably similar in all three cases (the attractive well is not visible on the scale of this figure) and so it seems again that the stationary functional gives reasonable physical properties even with very poor input densities.

D. The HHe molecule

H_2 and He_2 are both homonuclear molecules and so it seemed sensible to look at HHe as a first example of a heteronuclear molecule.

For the H atom, the spherical input density chosen was the exact atomic density and the basis set was the same as for the H_2 calculations. For the He atom, the input density was

$$n(r)=\frac{2(1.55)^3}{\pi}e^{-2(1.55)r} ,$$

which, of the two densities used in the He_2 calculation, was the one which gave an atomic energy closest to Phillipson's value. The He-atom basis set was just one s orbital,

$$\phi(r)=2(1.55)^{3/2}e^{-(1.55)r}Y_{00}(r) ,$$

exactly as in the He_2 calculation.

The results are shown in Fig. 4 and, for comparison, the results of the Hartree-Fock calculation of Fischer and Kemmey[43] are also plotted. The sum of the total energies for the separated atoms differs by a few percent from the Hartree-Fock value, but the interaction energies agree very well. As before, the calculated forces between the atoms are sensible even when the input densities are not very accurate.

E. The germanium dimer

As an example of a rather more complicated molecule, and to test the implementation of the stationary functional with a pseudopotential, it was decided to look at Ge_2. The advantage of dealing with germanium is that it can be well represented using a local pseudopotential of the Starkloff-Joannopoulos[44] (SF) form with parameters (see SJ for definitions) $r_c=1.052a_0$, $\lambda=18.5a_0^{-1}$, and $Z_v=4$. Although this pseudopotential is only approximately norm conserving, it produces excellent results for the solid, and calculations of the bulk modulus and lattice parameter give values very close to the experimental ones.[45] As far as we know, it has not yet been tested on molecules. In fact, a very similar but even simpler pseudopotential of the form

$$V_{\text{ps}}(r)=-\frac{Z_v}{r}\left[\frac{1-e^{-\lambda r}}{1+e^{\lambda(r_c-r)}}\right] ,$$

with $r_c=1.05a_0$, $\lambda=18.0a_0^{-1}$, and $Z_v=4$, gives almost identical results, at least for the dimer, and was used in the calculations described here.

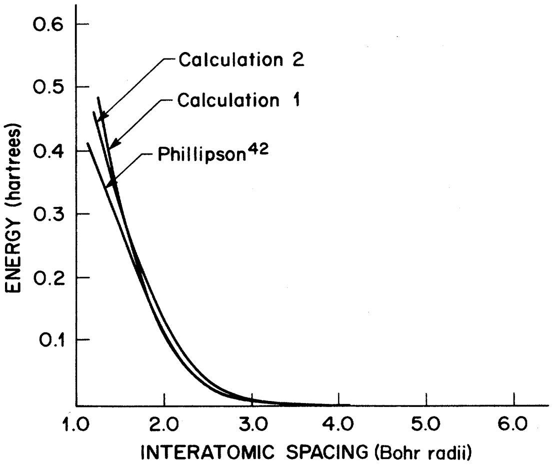

FIG. 3. Total energy of helium dimer (relative to the energy of the separated atoms) as a function of interatomic spacing.

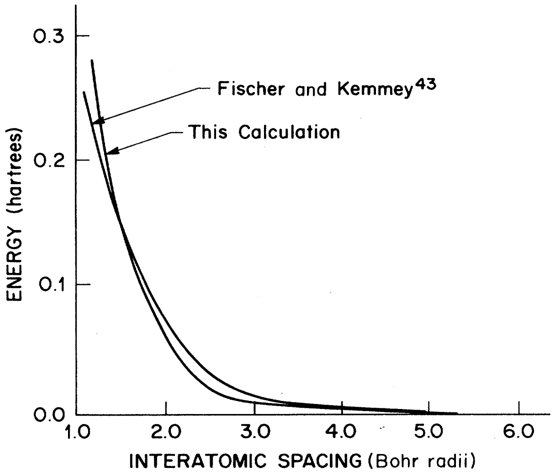

FIG. 4. Total energy HHe molecule as a function of interatomic spacing relative to the energy of the separated atoms.

It was decided to use superposed pseudoatomic LDA densities as input densities and these were calculated using a self-consistent atomic program written by Froyen.[46] The atomic program was not used in a spin-polarized mode and so the results are not very accurate for the free atoms; however, there is no reason to believe that superposing exact atomic densities is, in general, a particularly good thing to do and so the small errors in the atomic calculation are not important. (Indeed, as crystalline Ge is not spin polarized, one could even argue that superposing these imaginary non-spin-polarized atomic densities should be better than superposing exact atomic densities for the bulk.) The atomic calculation also produced 4s and 4p pseudoatomic wave functions and these were used as the s and p parts of the basis set. For the d orbitals, when included, we used the simple analytic form

$$\phi_d(r)\propto r^2e^{-\alpha r} ,$$

with $\alpha=1.175$. This value of α was found by Smith[47] to produce the best fit to the Cohen-Bergstrasser empirical pseudopotential calculations of the band structure of crystalline germanium. Smith also used s and p orbitals which were simple exponentials and a different pseudopotential from the one used here and so there is no reason to think that his value of α is in any way an optimum one in this case. However, it was found that small changes in α made little difference to the results of the dimer calculation and so Smith's value seemed as good a one to take as any.

Experimenting with the s and p functions by scaling them,

$$\phi_s(r)\rightarrow\beta_s^{3/2}\phi_s(\beta_s r), \quad \phi_p(r)\rightarrow\beta_p^{3/2}\phi_p(\beta_p r) ,$$

showed that $\beta_s=1$ and $\beta_p=1$ are very close to the (variationally) best values for the dimer. It is not worthwhile using swollen or contracted atomic orbitals as has been observed in some other cases. The effects of scaling the atomic input density,

$$n(r)\rightarrow\gamma^3 n(\gamma r) ,$$

were also examined. Table II shows the total dimer energy when the interatomic spacing was $4.8a_0$ (which is close to the equilibrium value) as a function of the parameter, γ, defining the input density. It can be seen that the stationary point is near $\gamma=1$ (atomic input densities) and at the total energy there is a *maximum*. This is a clear reminder that the functional being used is stationary but not variational.

The atomic program used to generate the density and s and p wave functions allows one to choose the electronic configuration of the atom. For an isolated atom, the ground state is $4s_2 4p_2$; in solid germanium, which a chemist would describe as sp_3 hybridized, one might expect something closer to $4s_1 4p_3$ (in fact, the configuration in the solid is somewhere between these two extremes). For the dimer calculations we tried a density and wave functions generated using the configuration $4s_{1.25}4p_{2.75}$ as well as those for the free-atomic configuration (in the investigations of scaling mentioned above the free-atomic configuration was used).

The results of these calculations are compared in Table III to the LMTO LSD calculations of Harris and Jones (HJ).[48] Unfortunately, HJ only quote values for the binding energy E_b, the equilibrium spacing R_e, and the vibrational frequency w_e, and so these were the only quantities we could compare. The inclusion of d functions was necessary to obtain a good interatomic spacing but had little effect on the binding energy or the vibrational frequency.

TABLE II. Total energy of a germanium dimer (interatomic spacing of 4.8 Bohr radii) as a function of γ, the parameter scaling the input density.

γ	Energy (hartrees)
0.75	−7.9119
0.80	−7.8182
0.85	−7.7482
0.90	−7.7015
0.95	−7.6734
1.00	−7.6635
1.05	−7.6700
1.10	−7.6903
1.15	−7.7319
1.20	−7.7899
1.25	−7.8578

TABLE III. Results of different calculations of the equilibrium interatomic spacing, R_e, the binding energy, E_b, and the vibrational frequency, w_e, for Ge_2.

	R_e/a_0	w_e (cm^{-1})	E_b (hartrees)
Harris and Jones[a]	4.75	240	0.0724
This work s and p orbitals (configuration $4s_2 4p_2$)	4.90	240	0.154
This work s, p, and d orbitals (configuration $4s_2 4p_2$)	4.73	240	0.141
This work s, p, and d orbitals (configuration $4s_{1.25}4p_{2.75}$)	4.65	280	0.162

[a] Reference 48.

The agreement is excellent except for the binding energies. These were calculated by subtracting the dimer energy from the energy of two isolated spin-zero germanium atoms and the problem, of course, is that a real germanium atom is spin polarized with a spin of 1. On average, three of the four valence electrons are up and only one is down, $n_+=3n/4$ and $n_-=n/4$. Using the same "exchange-only" approximation as was used for the H atom, this suggests that the xc energy in a real Ge atom will be approximately

$$\frac{(\frac{3}{4})^{4/3}+(\frac{1}{4})^{4/3}}{(\frac{1}{2})^{4/3}+(\frac{1}{2})^{4/3}}=1.057$$

times that in a spinless one. The unpolarized xc energy was −0.96 hartree per atom (for the $4s_2 4p_2$ configuration) and so the total energy of a spin-polarized atom should be roughly 0.055 hartree lower than that of a spinless one. Since there are two atoms in the dimer, this is about the right size to explain the discrepancy in the binding energy.

Apart from these familiar and unimportant problems, the results obtained using the stationary functional and pseudoatomic input densities are not appreciably different from those obtained by Harris and Jones with the much more accurate self-consistent LMTO method.

In Fig. 5, we have shown that part of the stationary functional (the double-counting terms evaluated at the input density plus the nuclear-nuclear Coulomb repulsion) which we claim gives the pair potential in the tight-binding energy expression. As can be seen, it is repulsive and roughly exponential in form—very similar to the pair potentials used in most semiempirical tight-binding schemes.

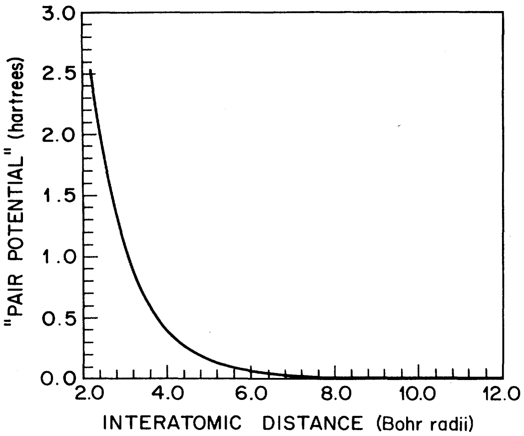

FIG. 5. Sum of double-counting energy and internuclear Coulomb energy for germanium dimer.

V. CONCLUSIONS

In this paper, we addressed some of the questions about the relation between the various parameters in non-self-consistent SETB calculations and self-consistent DFT. We saw that these questions fall into two categories: general questions about the forms of the TB energy expression, Eq. (14), and the corresponding non-self-consistent independent-electron equation, (15); and rather more specific and technical questions concerning the reduction of Eq. (15) to a simple secular equation with short-range and transferable matrix elements. With the use of a stationary principle in DFT, we showed how the questions in the first group can be understood and gave a prescription for calculating all the quantities appearing in Eqs. (14) and (15) (including the pair potential) direct from DFT.

Since this prescription produces a one-electron potential which is strictly local (in that the potential at any point depends only on the positions of the atoms near that point) it is in principle a straightforward matter to obtain the secular equation for any arrangement of atoms. All that is necessary is to choose a basis set and evaluate a set of three-dimensional integrals. This was done for a selection of dimers and the stationary functional performed excellently, producing good results even for very poor input densities. In addition, it turned out that very simple basis sets, consisting of just a few atomiclike orbitals on each atom, were adequate, which is another encouraging sign.

Assuming that the use of the stationary functional with simple input densities proves successful in solids (and some recent work by Polatoglou and Methfessel[32] suggests that it will) and that similarly unsophisticated basis sets will suffice, there seems no reason why one should not be able to derive equally accurate and reliable tight-binding parameters for real solid-state applications. The only obvious remaining hurdle is the evaluation of the three- (and four-) center integrals necessary to find the Hamiltonian matrix in a solid. These involve the potential near one atom and the basis functions from two others and are numerous and difficult both to calculate and to parametrize (so that they can be stored for future use). Unfortunately, they are not negligible as became clear when we attempted to calculate the properties of crystalline germanium without including them. Although the total (valence) energy at the experimental equilibrium spacing was accurate to considerably better than 1% (with *s*, *p*, and *d* orbitals as used in the dimer calculations and the same pseudopotential), we did not get a good lattice parameter and the calculated bulk modulus was far from the experimental one. We are convinced that these problems are mainly due to the neglect of the three-center integrals and have some ideas as to how to improve matters, but the topic is sufficiently complicated that we leave it for a future paper.

ACKNOWLEDGMENTS

The work was supported by the National Science Foundation under Grant No. DMR-87-12346. One of us (W.M.C.F.) would also like to acknowledge the support of Pembroke College and The Drapers' Company.

---

[*] Present address: AT&T Bell Laboratories, 600 Mountain Avenue, Murray Hill, NJ 07974-2070.

[†] Present address: Theory of Condensed Matter Group, Cavendish Laboratory, University of Cambridge, Madingley Road, Cambridge CB3 0HE, United Kingdom.

[1] See, for instance, W. A. Harrison, *Pseudopotentials in the Theory of Metals* (Benjamin, New York, 1966).

[2] J. K. Nørskov and N. D. Lang, Phys. Rev. B **21**, 2131 (1980); M. J. Stott and E. Zaremba, *ibid.* **22**, 1564 (1980).

[3] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964); W. Kohn and L. J. Sham, *ibid.* **140**, A1133 (1965); for a review, see C. J. Callaway and N. H. March in *Solid State Physics,* edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic, New York, 1984), Vol. 38.

[4] R. G. Gordon and Y. S. Kim, J. Chem. Phys. **56**, 3122 (1972).

[5] R. Biswas and D. Hamann, Phys. Rev. Lett. **55**, 2001 (1985).

[6] R. Car and M. Parinello, Phys. Rev. Lett. **55**, 2471 (1985); M. C. Payne, P. D. Bristowe, and J. D. Joannopoulos, Phys. Rev. Lett. **58**, 1348 (1987).

[7] The articles by V. Heine, R. Haydock, D. Bullett, and M. J. Kelly, all in *Solid State Physics,* edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic, New York, 1980), Vol. 35, give a fairly complete summary of the uses of the semiempirical tight-binding method until 1980. For examples of some more recent work see S. Glanville, A. T. Paxton, and M. W. Finnis, J. Phys. F **18**, 693 (1988); A. P. Sutton, M. W. Finnis, D. G. Pettifor, and Y Ohta, J. Phys. C **21**, 35 (1988); W. A. Harrison, Phys. Rev. B **34**, 2787 (1986).

[8] D. J. Chadi, Phys. Rev. Lett. **43**, 43 (1979).

[9] M. T. Yin and M. L. Cohen, Phys. Rev. B **24**, 2303 (1981); K.

C. Pandey, in *Proceedings of the 17th International Conference on the Physics of Semiconductors, San Francisco, 1984*, edited by D. J. Chadi and W. A. Harrison (Springer, New York, 1985), p. 55.

[10] R. M. Tromp, R. J. Hamers, and J. E. Demuth, Phys. Rev. Lett. **55**, 1303 (1985); R. J. Hamers, R. M. Tromp, and J. E. Demuth, Surf. Sci. **181**, 346 (1987); P. E. Wierenga, J. A. Kubby, and J. E. Griffith, Phys. Rev. Lett. **59**, 2169 (1987); J. A. Kubby, J. E. Griffith, R. S. Becker, and J. S. Vickers, Phys. Rev. B **36**, 6079 (1987); M. Needels, M. C. Payne, and J. D. Joannopoulos, Phys. Rev. Lett. **58**, 1765 (1987).

[11] M. Hashimoto, Acta Metall. **32**, 13 (1984).

[12] C. M. Sayers, Philos. Mag. **50**, 635 (1984).

[13] L. Goodwin, R. J. Needs, and V. Heine, Phys. Rev. Lett. **60**, 2050 (1988); and Erratum, Phys. Rev. Lett. **61**, 133 (1988).

[14] R. Haydock, J. Phys. C **14**, 3807 (1981).

[15] J. Harris, Phys. Rev. B **31**, 1770 (1985).

[16] W. M. C. Foulkes, Pre-Ph.D. thesis, 1985, and Ph.D. thesis, University of Cambridge, 1988.

[17] W. M. C. Foulkes and R. Haydock, J. Phys. C **19**, 6573 (1986).

[18] M. Levy, Phys. Rev. A **26**, 1200 (1982).

[19] L. Hedin and B. I. Lundqvist, J. Phys. C **4**, 2064 (1971).

[20] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).

[21] J. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).

[22] See, for example, N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Holt-Saunders, Tokyo, 1976), p. 187.

[23] J. C. Slater and G. F. Koster, Phys. Rev. **94**, 1498 (1954).

[24] A. R. Mackintosh and O. K. Andersen, in *Electrons at the Fermi Surface*, edited by M. Springford (Cambridge University Press, Cambridge, 1980), p. 149.

[25] W. A. Harrison, *Electronic Structure and the Properties of Solids* (Freeman, San Francisco, 1980).

[26] See the article by D. Bullett in Ref. 7.

[27] O. K. Andersen, O. Jepsen, and M. Sob, in *Electronic Band Structure and Its Applications*, edited by M. Yussouf (Springer, New York, 1987).

[28] H. Wendel and R. M. Martin, Phys. Rev. B **19**, 5251 (1979).

[29] U. von Barth and C. D. Gelatt, Phys. Rev. B **21**, 2222 (1980).

[30] G. B. Bachelet, D. R. Hamann, and M. Schlüter, Phys. Rev. B **26**, 4199 (1982); G. P. Kerker, J. Phys. C **13**, 189 (1980).

[31] S. G. Louie, S. Froyen, and M. Cohen, Phys. Rev. B **26**, 1738 (1982).

[32] H. M. Polatoglou and M. Methfessel, Phys. Rev. B **37**, 10 403 (1988).

[33] C. M. M. Nex, M.Sc. thesis, Sussex University, 1967.

[34] O. Gunnarsson, B. I. Lundqvist, and J. W. Wilkins, Phys. Rev. B **10**, 1319 (1974).

[35] R. F. W. Bader, in *The Force Concept in Chemistry*, edited by B. M. Deb (Van Nostrand/Reinhold, New York, 1981), p. 102.

[36] W. Kolos and C. J. Roothaan, Rev. Mod. Phys. **32**, 219 (1960).

[37] O. Gunnarsson and P. Johansson, Int. J. Quantum Chem. **10**, 307 (1976).

[38] G. S. Painter, Phys. Rev. B **24**, 4264 (1981).

[39] O. Gunnarsson and B. I. Lundqvist, Phys. Rev. B **13**, 4274 (1976).

[40] S. H. Vosko, L. Wilk, and M. Nusair, Can. J. Phys. **58**, 1200 (1980).

[41] F. Herman and S. Skillman, *Atomic Structure Calculations* (Prentice Hall, Englewood Cliffs, 1963).

[42] P. E. Phillipson, Phys. Rev. **125**, 1981 (1962).

[43] C. R. Fischer and P. J. Kemmey, J. Chem. Phys. **53**, 50 (1970).

[44] T. Starkloff and J. D. Joannopoulos, Phys. Rev. B **16**, 5212 (1977).

[45] M. C. Payne (private communication).

[46] We thank Sverre Froyen and R. J. Needs for allowing us to use this program.

[47] P. V. Smith, Phys. Status Solidi B **116**, 101 (1983).

[48] J. Harris and R. O. Jones, Phys. Rev. A **18**, 2159 (1978).