

Contemporary pseudopotentials—Simple reliability criteria

Alex Zunger

Citation: [Journal of Vacuum Science & Technology](#) **16**, 1337 (1979); doi: 10.1116/1.570153

View online: <http://dx.doi.org/10.1116/1.570153>

View Table of Contents: <http://scitation.aip.org/content/avs/journal/jvst/16/5?ver=pdfcov>

Published by the [AVS: Science & Technology of Materials, Interfaces, and Processing](#)

Articles you may be interested in

[Nonlocal pseudopotential calculation of the surface tension of simple liquid metals](#)

J. Chem. Phys. **86**, 2095 (1987); 10.1063/1.452159

[Simple criteria for the absence of the beam-Weibel instability](#)

Phys. Fluids **24**, 1818 (1981); 10.1063/1.863262

[Solitons and simple pseudopotentials](#)

J. Math. Phys. **17**, 756 (1976); 10.1063/1.522974

[Comparison of A-Weighted Auditory Risk Criteria With Contemporary Octave-Band Estimates of Auditory Damage](#)

J. Acoust. Soc. Am. **50**, 110 (1971); 10.1121/1.1977484

[Pseudopotential Calculation of the Phonon Frequencies of Simple Metals](#)

J. Chem. Phys. **54**, 435 (1971); 10.1063/1.1674629

ADVERTISEMENT



 Advance your technology or engineering career using the **AVS Career Center**, with **hundreds of exciting jobs** listed each month!

<http://careers.avs.org>



Contemporary pseudopotentials—Simple reliability criteria

Alex Zunger

Solar Energy Research Institute, Photovoltaics Branch, Golden, Colorado 80401

(Received 21 March 1979; accepted 24 May 1979)

Simple tests are presented that gauge the accuracy of some of the pseudopotentials currently used in the literature to calculate within the density-functional approach the electronic properties of molecules, solids and surfaces. These include the local “soft-core” and “hard-core” potentials as well as the nonlocal first principles (“hard-core”) potentials. A discussion of common misconceptions regarding pseudopotentials is included.

PACS numbers: 71.10. + x, 73.20. – r, 31.70. – f

I. INTRODUCTION

Recently, a number of questions have been raised pertaining to the reliability of the contemporary pseudopotential approaches in describing the electronic structure of molecules, solids and surfaces.^{1–6} In this paper we present a discussion of the underlying problems and suggest a number of elementary tests for gauging in a simple manner the quality of various pseudopotentials. We find that these tests tend to add strong support to the recently developed first principles nonlocal density functional pseudopotentials.^{5b} In the second part of the paper we comment on a number of common misconceptions regarding pseudopotentials.

II. CONTEMPORARY PSEUDOPOTENTIALS

Here we briefly define the approaches underlying a number of contemporary pseudopotentials used extensively in variational electronic structure calculations (i.e., non-perturbative approaches). We first establish the necessary terminology.

A. Terminology

The pseudopotential approach attempts to replace an all-electron (i.e., core and valence wavefunctions treated on the same footing) eigenvalue problem:

$$\left\{ -\frac{1}{2} \nabla^2 + V_{c,v}[\rho(\mathbf{r}, \mathbf{r}')] \right\} \psi_j^c(\mathbf{r}) = \epsilon_j \psi_j^c(\mathbf{r}), \quad (1)$$

with its associated core (*c*) and valence (*v*) total potential $V_{c,v}[\rho(\mathbf{r}, \mathbf{r}')]$ and wavefunctions $\psi_j^c(\mathbf{r})$, by a simpler equation:

$$\left\{ -\frac{1}{2} \nabla^2 + W_{ps}(\mathbf{r}) + V_v[n(\mathbf{r}, \mathbf{r}')] \right\} \chi_j(\mathbf{r}) = \lambda_j \chi_j(\mathbf{r}), \quad (2)$$

which applies to the valence subspace alone. The total potential $V_{c,v}$ includes electron-nuclear, interelectronic Coulomb, as well as exchange (and possibly correlation) terms. These interactions can be modeled by a number of possible electronic structure approaches such as Hartree-Fock, density functional, etc. Here $V_v[n(\mathbf{r}, \mathbf{r}')]$ has the same functional form as $V_{c,v}[\rho(\mathbf{r}, \mathbf{r}')]$ but is a function of the valence pseudocharge density alone:

$$n(\mathbf{r}, \mathbf{r}') = \sum_j^v \chi_j^*(\mathbf{r}) \chi_j(\mathbf{r}'), \quad (3)$$

whereas $\rho(\mathbf{r}, \mathbf{r}')$ is the all-electron charge density:

$$\rho(\mathbf{r}, \mathbf{r}') = \sum_j^{c+v} \psi_j^*(\mathbf{r}) \psi_j(\mathbf{r}') \equiv \rho_c(\mathbf{r}, \mathbf{r}') + \rho_v(\mathbf{r}, \mathbf{r}'). \quad (4)$$

The all-electron density has a core part (ρ_c) and a valence part (ρ_v). V_v is referred to as “screening” or “valence field” and can be constructed like $V_{c,v}$ according to a number of microscopic approaches (e.g., Hartree-Fock, density functional, $X\alpha$, free-electron dielectric screening, etc.). It includes an interelectronic Coulomb part $V_{\text{coul}}[n]$ and an exchange (and possibly correlation) part $V_{xc}[n]$. $W_{ps}(\mathbf{r})$ is the total pseudopotential. It is often presented as a sum over all atoms in the system (with position vectors \mathbf{R}_α) and all angular momenta (*l*), of the “atomic pseudopotential $V_{ps}^{(l)}(\mathbf{r})$ (sometimes referred to as “ionic potential” or “core potential”):

$$W_{ps}(\mathbf{r}) = \sum_\alpha \sum_l V_{ps}^{(l)}(\mathbf{r} - \mathbf{R}_\alpha) \hat{P}_{l,\alpha}, \quad (5)$$

where $\hat{P}_{l,\alpha}$ is the *l*th angular momentum projector with respect to the α th center.

An atomic pseudopotential that can be expressed as angular momentum dependent (and hence acts differently on different angular momentum components of the wavefunction) is said to be “nonlocal.” It is sometimes approximated by an average *l*-independent form, referred to as a “local atomic pseudopotential” $V_{ps}(\mathbf{r})$.

The total potential seen by an electron in the pseudo-system,

$$V_{\text{eff}}(\mathbf{r}) = W_{ps}(\mathbf{r}) + V_v[n(\mathbf{r}, \mathbf{r}')], \quad (6)$$

is referred to as the “effective potential.” In the special case of a single pseudo atom (or ion) one has an “atomic (ionic) effective potential” (or the “screened atomic [ionic] potential”):

$$V_{\text{eff},a}^{(l)}(\mathbf{r}) = V_{ps}^{(l)}(\mathbf{r}) + V_v[n(\mathbf{r}, \mathbf{r}')], \quad (7)$$

where V_v represents here atomic screening.

To the extent that the construction of the appropriate pseudopotentials $V_{ps}^{(l)}(\mathbf{r})$ can be made simple, the solution of the pseudopotential single-particle Eq. (2) for arbitrary systems is easier than the solution of the all-electron problem, Eq. (1), as a smaller number of “reactive” electrons are being

considered and the pseudo wavefunctions $\chi_j(\mathbf{r})$ can be made to be spatially smooth and hence computationally simpler. Once $V_{ps}^{(l)}(\mathbf{r})$ is calculated for some simple model system (e.g., atom, ion), it is assumed to be transferable to other systems containing the same cores ("pseudopotential frozen core approximation") and used for a self-consistent solution of Eqs. (2), (3), and (5).

As the pseudopotential Hamiltonian $-1/2\nabla^2 + V_{\text{eff}}$ permits (but does not require) nonorthogonal solutions $\chi_j(\mathbf{r})$, it has been a common practice to design $\chi_j(\mathbf{r})$ as nodeless and smooth in the core region. In doing so, one gives up from the outset the possibility of inferring core properties (e.g., Fermi contact interactions, core photoemission spectra, nuclear-quadrupole coupling constants, etc.) from $\chi_j(\mathbf{r})$, unless a special construction of the latter is undertaken. This is a distinct disadvantage of the pseudopotential approach over the all-electron one. In judging the quality of pseudopotentials, we hence limit the discussion to the degree V_{eff} can produce accurate "valence"-related constructs. In this paper I consider a given pseudopotential approach as successful if it simplifies the electronic structure problem, yields an accurate description of valence properties over a desired range of energies and chemical variations (where the properties and the relevant accuracy are problem-dependent and are to be defined below), and carries over from the all-electron representation the chemical regularities underlying the response of the outer valence electrons.

There is an entire class of pseudopotential approaches that do not attempt the calculation of the pseudopotential for reference systems but rather that pertaining directly to the polyatomic system of interest, (e.g., Ref. 7). We do not discuss these approaches here as they are not normally applied to the electronic structure of complex solids or surfaces.

B. Self-consistent versus empirical potentials

In the "empirical pseudopotential"^{8,9} (EPM) approach to the electron structure of solids one avoids completely the microscopic specification of both the pseudopotential $V_{ps}^{(l)}(\mathbf{r})$ and the screening $V_e(\mathbf{r})$. Instead, the atomic effective potential $V_{\text{eff},a}$ is represented by a set of values $V_{\text{eff},a}(\mathbf{G}_i)$ taken at the first few reciprocal vectors \mathbf{G}_i of a given lattice (i typically ranging from 3 to 5). These are used as disposable parameters adjusted to reproduce a chosen set of experimental observables relating to the energy eigenvalues λ_j (e.g., the lowest interband transitions in semiconductors, the Fermi surface of metals, etc.). The quantity treated as system-invariant (or "transferable") is, hence, the *screened* ionic potential [(Eq. (7))]. As no constraints are placed on the corresponding crystal wavefunctions, and since the *relative* dispersion of the energy bands in the valence region of a solid is insensitive to the large momentum-transfer scattering events $V_{\text{eff}}(q \gg 2k_F)$ (which sample the core region), it has been possible,^{8,9} to obtain good fits to the interband transition energies using small momentum cutoff values in $V_{\text{eff},a}(\mathbf{G})$ (i.e., implying "weak" or "soft" potentials).

Self-consistency (or the response of the electronic system to the changed chemical environment relative to the zero-order potential) is outside the scope of this approach as this form of the effective potential is unrelated to the variational

wavefunctions. In more recent calculations,^{10,11} an empirical "nonlocal" component has been added to $V_{\text{eff},a}(\mathbf{r})$ (in the form of a parametrized square well or Gaussian). Reparametrization in this form has led to an overall improvement of the agreement with the observed optical and photoemission data as well as with the experimentally synthesized charge density. Many applications of this approach to covalently bonded semiconductors^{8,9} indicate an excellent reproduction of the observed optical transitions and at the same time a chemically reasonable charge density is obtained. Because of an inherent lack of a self-consistency mechanism, this approach has not been applied to systems that differ substantially in their chemical environment (e.g., molecules, surfaces and atoms) from the reference fitted system (e.g., bulk solid). Hence, in what follows we will not discuss this approach.

Self-consistent pseudopotential approaches are based on the partitioning of the electronic interactions represented in Eq. (6). The screening is calculated self-consistently from the density $n(\mathbf{r})$ of the *system of interest*, while the pseudopotential is fixed from the outset for a *reference system*. Only $V_{ps}^{(l)}(\mathbf{r})$ is considered as the system-invariant quantity. We proceed in defining the various approaches used to calculate $V_{ps}^{(l)}(\mathbf{r})$.

C. Semiempirical pseudopotentials

In the semiempirical pseudopotential (SEMP) method one usually assumes an *ansatz* functional form for $V_{ps}(\mathbf{r})$ or $V_{ps}(\mathbf{q})$ with a number of free parameters that are adjusted to reproduce either the observed free-ion term values¹² or the bulk band structure.¹³ No direct constraints,¹³ or at best very weak constraints, are placed on the resulting wavefunctions (e.g., matching the point of maximum of the pseudo and all-electron atomic orbitals). As this process of fitting the low-energy eigenvalue spectra of a pseudopotential single-particle equation can be carried out with arbitrary forms of $V_{ps}(\mathbf{r})$ in the core region, a soft-core form (finite and smooth in the core) has usually been chosen.¹⁴ This choice is guided by the need to have a momentum-space converged potential when a plane wave basis is used for expanding the crystal wavefunctions.¹² As empirical fitting is involved, the role of correlation effects on the electronic structure cannot be assessed.

The majority of the self-consistent surface calculations to date have been performed with such potentials (e.g., Ref. 15 and references therein). The two most frequently used forms for silicon are the Berkeley potential¹² and the Appelbaum-Hamann¹³ potential. Both use a local approximation and yield very similar results. The predictions of those potentials for the electronic structure of a Si_2 diatomic molecule^{1,3} and bulk Si^4 were recently contrasted with the corresponding all-electron results. Differences in the molecular cohesive energy, orbital charge densities and band structure were identified. The reader is referred to these papers for judging the significance of these discrepancies.

D. Model pseudopotentials

In the model pseudopotential approach one assumes a parametrized form for $V_{ps}^{(l)}(\mathbf{r})$ and adjusts these parameters

to fit some *theoretical* constructs of an all-electron calculation on a reference system. A recent example of a model potential is that of Harris and Jones² (for Si) in which the orbital energies as well as the wavefunctions of a density-functional Si atom were fitted within a postulated local form $V_{ps}(\mathbf{r})$. A good fit has been achieved within a “hard-core” (strongly repulsive at small- \mathbf{r}) form. A similar approach has been successfully applied by Louie to a Pd surface.¹⁶ As this approach is completely nonempirical, the role of correlation effects can be assessed. Recent tests on the Si₂ molecule² and bulk Si⁴ indicate good agreement with the all-electron results.

E. First principles pseudopotentials

The first principles pseudopotential approach calculates $V_{ps}^{(l)}(\mathbf{r})$ from a given microscopic electronic structure theory. The form of $V_{ps}^{(l)}(\mathbf{r})$ is not assumed as an ansatz^{2,10–13} as there is no simple physical rule that dictates its shape at small \mathbf{r} . No empirical data (atomic or solid state) are used. The form of the corresponding wavefunctions is not allowed to be implicitly determined by an arbitrary choice of the functional form of $V_{ps}^{(l)}(\mathbf{r})$. Instead, the pseudopotential is constructed to *explicitly* satisfy certain constraints on the wavefunctions. First-principles pseudopotentials in the Hartree–Fock (HF) and correlated HF approaches were constructed by Goddard *et al.*,¹⁷ and Kahn *et al.*,¹⁶ and applied to a large number of molecular^{19,20} and cluster²¹ calculations. First-principles density-functional pseudopotentials have been developed by Zunger *et al.*,²² and applied to molecules,^{3,23} bulk semiconductors,^{5a} electronic and cohesive properties of bulk transition metals,²² the relaxed GaAs(110) surface²⁵ and to the prediction of the crystal structure of binary AB compounds.²⁶

Two basic approaches are possible here: the first attempts to construct the “best” pseudopotential possible within the orbital space characteristic of either the Hartree–Fock (Hartree–Fock pseudopotentials¹⁹) or the density-functional (density functional pseudopotential,²² or DF). This means that the corresponding pseudo orbitals be represented as a simple unitary rotation of the corresponding core and valence all-electron orbitals (see below). The second approach²⁷ (trans-Hartree–Fock or trans-density-functional, TDF) is a modification of the first and allows for certain specified wavefunction components to lie outside the corresponding orbital spaces. This approach still attempts to describe the underlying all-electron results (and not anything better) using a more general representation for the basis. The detailed description of these methods is presented elsewhere.²⁷ Here we describe in simple terms the DF and the TDF approaches (as applied below to Si).

1. The density-functional (DF) pseudopotential

We first set a variationally derived single-particle equation for a pseudo-atom having only valence electrons that interact via Coulomb ($V_{coul}[n(\mathbf{r})]$) and the density-functional exchange-correlation $\{V_{xc}[n(\mathbf{r})]\}$ potentials and with a yet unspecified external potential $V_{ps}^{(l)}(\mathbf{r})$:

$$\left\{ -\frac{1}{2} \nabla^2 + V_{ps}^{(l)}(\mathbf{r}) + V_{coul}[n(\mathbf{r})] + V_{xc}[n(\mathbf{r})] \right\} \chi_{nl}(\mathbf{r}) = \lambda_{nl} \chi_{nl}(\mathbf{r}). \quad (8)$$

We require that the normalized pseudo-orbitals χ_{nl} span the same space characteristics of the all-electron density function problem:

$$\chi_{nl}(\mathbf{r}) = \sum_{n'}^{c+v} C_{n,n'}^{(l)} \psi_{n'l}(\mathbf{r})$$

$$n(\mathbf{r}) = \sum_{nl} \chi_{nl}^*(\mathbf{r}) \chi_{nl}(\mathbf{r}) \quad (9)$$

where the sum extends on all (known) core and valence orbitals of symmetry l . The total valence energy of the system is unchanged by such a unitary mixing. In addition, one can recover at the end of the calculation the true valence orbitals simply by core orthogonalization.

One then imposes certain constraints on the energy eigenvalues λ_{nl} and the orbitals $\chi_{nl}(\mathbf{r})$ (see below) at a given electronic configuration (g) of the atom and solves Eqs. (8) and (9) for the potential $V_{ps}^{(l)}(\mathbf{r})$ that satisfies these Lagrange-multiplier constraints. One then assumes that this static potential $V_{ps}^{(l)}(\mathbf{r})$ that replaces the dynamic effects of the core electrons for the reference configuration g , will continue to be a valid approximation for configurations other than g and for the same core in different chemical environments (i.e., the pseudopotential frozen core approximation).

The constraints imposed are:²² (i) the valence orbital energy spectrum λ_{nl} of Eq. (8) be equal to that of the all-electron system, (ϵ_{nl} of Eq. 1) in the same configuration, (ii) $\chi_{nl}(\mathbf{r})$ be normalized, (iii) $\chi_{nl}(\mathbf{r})$ be nodeless for each of the low-symmetry states (as required of any ground state central-field problem), (iv) the coefficients $\{C_{n,n'}^{(l)}\}$ be determined such that the pseudo orbitals have the maximum similarity (possible under the previous constraints) to the corresponding “true” orbitals $\psi_{nl}(\mathbf{r})$ in the chemically relevant tail region. This is accomplished by minimizing the core projection of $\chi_{nl}(\mathbf{r})$ for all the valence orbitals. The most effective minimization of a smooth $\chi_{nl}(\mathbf{r})$ [of the form of Eq. (9)] in the core requires that the orbital vanishes at the origin. This produces pseudo orbitals with a very low amplitude in the core region [reflecting their closeness to the true orbitals in the valence regions, cf. Eq. (9)] and hence the necessity for core orthogonalization is drastically reduced.

Such a pseudopotential can be written in a simple closed form. Making use of the fact that the $\psi_{nl}(\mathbf{r})$ in Eq. (9) satisfy the all-electron Eq. (1), one obtains:

$$V_{ps}^{(l)}(\mathbf{r}) = \left\{ U_l(\mathbf{r}) - \frac{Z_v}{r} \right\} + \left\{ -\frac{Z_c}{r} + V_{coul}[\rho_c(\mathbf{r})] + V_{xc}[\rho_c(\mathbf{r})] \right\} + \{V_{xc}[\rho_c(\mathbf{r}) + \rho_v(\mathbf{r})] - V_{xc}[\rho_c(\mathbf{r})] - V_{xc}[\rho_v(\mathbf{r})]\} + \{V_{coul}[\rho_v(\mathbf{r})] - V_{coul}[n(\mathbf{r})]\} + \{V_{xc}[\rho_v(\mathbf{r})] - V_{xc}[n(\mathbf{r})]\}. \quad (10)$$

Here Z_c and Z_v denote the valence and core charges, $\rho_c(\mathbf{r})$ and $\rho_v(\mathbf{r})$ are the corresponding charge densities from the all-electron calculation [Eq. (4)], and the repulsive “Pauli potential” $U_l(\mathbf{r})$ is:

$$U_l(\mathbf{r}) = \frac{\sum_{n'} C_{n,n'}^{(l)} [\epsilon_{nl}^v - \epsilon_{n'l}^{c,v}] \psi_{n'l}(\mathbf{r})}{\sum_{n'} C_{n,n'}^{(l)} \psi_{n'l}(\mathbf{r})}. \quad (11)$$

The pseudopotential in Eq. (10) has a very transparent physical interpretation. The Pauli potential $U_l(\mathbf{r})$ contains all the nonlocality, while all other terms are local. For valence states with matching symmetries in the core (e.g., Si 3s, 3p), $U_l(\mathbf{r})$ is strongly repulsive (with a small- r behavior of C_l/r^2) and replaces the loss of the orbital kinetic energy through the elimination of the nodal character in χ_{nl} . For valence states with no matching symmetries in the core (e.g., carbon 2p, silicon 3d, etc.), the valence orbital $\psi_{nl}(\mathbf{r})$ is already nodeless in the all-electron representation (hence $C_{n,n'} = \delta_{n,n'}$) and $U_l(\mathbf{r})$ is identically zero by Eq. (11) and no pseudopotential kinetic energy cancellation [i.e., partial cancellation of $U_l(\mathbf{r})$ by the core-attraction $-Z_v/r$] is taking place.

The second term in brackets in Eq. (10) is simply the potential set by the core electrons in the atom. The third term represents the nonlinearity of the exchange-correlation potential with respect to the interfering core and valence charge densities. The fourth and fifth terms are, respectively, the Coulomb and exchange-correlation orthogonality hole potentials and emerge from the elimination of the structure characteristics of $\rho_v(\mathbf{r})$ in the pseudo density $n(\mathbf{r})$ due to the cancellation of the nodes in $\chi_{nl}(\mathbf{r})$.

The first principles DF pseudopotentials are nonlocal and can be characterized as hard-core for those l -components present in the core. This hard-core behavior is not arbitrarily imposed in the construction of $V_{ps}^{(l)}(\mathbf{r})$ but is rather a consequence of the maximum similarity constraint imposed on the wavefunctions with a minimum amplitude of $\chi(\mathbf{r})$ in the core. As the potential has a microscopically transparent form, both its successes and failures in describing the electronic structure of various systems can be analyzed in terms of the constructs of the underlying DF theory.

Note that the repulsiveness of the DF pseudopotentials in the core region can be artificially reduced by including in the pseudo wavefunctions $\chi_{nl}(\mathbf{r})$ [Eq. (9)] more core components $\psi_{nl}^c(\mathbf{r})$ than are needed just to eliminate the nodes in the pseudo wavefunctions. This results in a higher amplitude of $\chi_{nl}(\mathbf{r})$ in the core region at the expense of a lesser similarity to the true orbitals $\psi_{nl}^v(\mathbf{r})$ in the valence region. The increase of $\chi_{nl}(\mathbf{r})$ in the core offsets the repulsiveness of the Pauli potential $U_l(\mathbf{r})$ [Equation (11)] resulting in a "weaker" pseudopotential and, ultimately, in a "soft-core" potential.²² Hence, there exists a continuous tradeoff between the reduced repulsiveness of $V_{ps}^{(l)}(\mathbf{r})$ in the core region and the variational quality of the wavefunctions in the chemically relevant tail region.

2. The Trans-density functional (TDF) pseudopotentials

Although the DF pseudopotential is constructed to yield orbitals with a maximum degree of similarity to the all-electron orbitals (possibly within this orbital space), there is an internal limit to the degree of similarity attainable. This is a consequence of the fact that $\{\psi_{nl}(\mathbf{r})\}$ forms an orthogonal space. Hence, for the Si 3s orbital Eq. (9) reads:

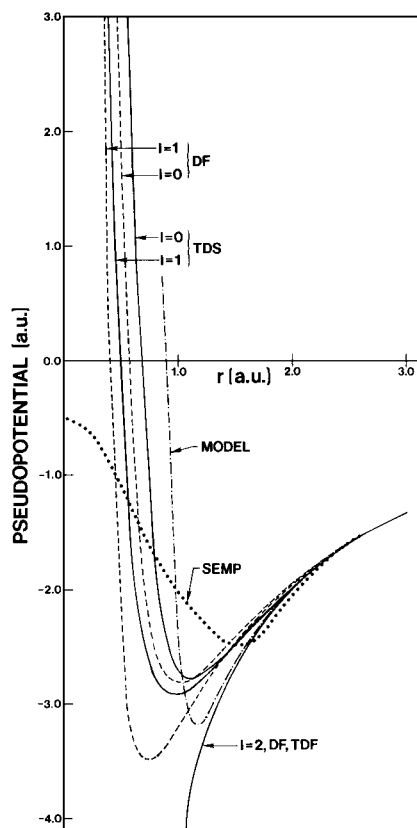


FIG. 1. Si pseudopotentials. TDF: nonlocal trans density functional;²⁷ DF: nonlocal density functional;²² MODEL: the local Harris and Jones² model potential; SEMP: the local semiempirical Berkeley¹² potential.

$$\chi_{3s}(\mathbf{r}) = C_{1s,3s}^{(o)} \psi_{1s}(\mathbf{r}) + C_{2s,3s}^{(o)} \psi_{2s}(\mathbf{r}) + C_{3s,3s}^{(o)} \psi_{3s}(\mathbf{r}) \quad (12)$$

and due to the orthogonality of $\{\psi_{nl}(\mathbf{r})\}$:

$$C_{1s,3s}^2 + C_{2s,3s}^2 + C_{3s,3s}^2 = 1 \quad (13)$$

Hence, $C_{3s,3s} < 1$, which implies a finite and constant difference between $\psi_{3s}(\mathbf{r})$ and $\chi_{3s}(\mathbf{r})$ in the tail region. This difference is explicitly minimized in the DF approach, but cannot be zeroed. As a result, typical orbital moments $\langle \chi | \mathbf{r}^p | \chi \rangle$ for $p = 1, 2$ can differ from the all-electron moments $\langle \psi | \mathbf{r}^p | \psi \rangle$ by 2% to 5%. Such a difference, although seemingly small, can introduce some distortions in the large- r behavior of the charge density.²⁸ If one desires to eliminate these differences and impose the maximum similarity constraint in a more stringent manner (e.g., reduce the differences to below 1%), an element of nonorthogonality has to be introduced into Eq. (9). This means that we have to go outside the space spanned by $\psi_{nl}(\mathbf{r})$, in some controlled fashion. As the tail deviation of $\chi_{nl}(\mathbf{r})$ from $\psi_{nl}(\mathbf{r})$ is known, however, one can simply achieve this goal by adding to $\chi_{nl}(\mathbf{r})$ in Eq. (9) a simple function $f_l(\mathbf{r})$ designed to correct the tail behavior (without affecting the small- r behavior). A full discussion of this approach is given elsewhere.²⁷ Here we simply note that using the choice:

$$f_l(\mathbf{r}) = A_l r^{l+2} e^{-\alpha_l r}, \quad (14)$$

and adjusting A_l and α_l to produce the correct tail matching yields the desired improvement. The corresponding TDF pseudopotential is very similar in shape to the DF potential (Fig. 1) except that it is shifted somewhat away from the core. It produces a close approximation to the valence part of the all-electron density matrix.

3. Difficulties with the DF and TDF pseudopotentials

The first principles DF and TDF pseudopotential suffer from the following problems:

(a) As fitting to the observed excitation energies is avoided, these pseudopotentials are unable to reproduce the physical factors which are not included in the chosen underlying all-electron problem. These may include relativistic, spin-orbit and certain classes of long-range correlation effects. The generalization of these potentials to include the first two effects is straightforward. Introduction of correlation effects is substantially more difficult.^{5b}

(b) The hard-core nature of these potentials makes them unsuitable for a simple plane-wave basis band structure model. Instead, they require a basis set which contains localized functions such as Gaussians, more akin to the conventional quantum chemical description of molecules and solids.

(c) As their construction involves the knowledge of the all-electron atomic energies and orbitals, the effort required increases with atomic number Z . In contrast, the construction of the empirical and semiempirical pseudopotentials involves an effort which is Z -independent once the relevant spectroscopic data are available.

III. SIMPLE CRITERIA FOR ATOMIC PSEUDOPOTENTIALS

In this section we consider simple tests for the reliability of the pseudopotential discussed in Sec. II. We would like to advocate the use of the electronic structure of the *atom* (and its neutral and ionized excited states) as a useful probe for the reliability of the potential. This is a well defined limit that can be solved numerically with almost arbitrary internal accuracy (typically 10^{-6} Hartrees) as no essential computational approximations are done to the potential or the orbitals. The simplicity of the system, and the widespread availability of efficient computer programs which accurately solve an atomic eigenvalue problem, can make these and similar tests easily reproducible by many nonspecialized practitioners. By means of exciting the atom we are able to probe fairly wide regions of wavefunction localizations and orbital energies, testing thereby the transferability of the pseudopotentials.

As various pseudopotential applications tend to emphasize different aspects of the electronic structure (e.g., small- r versus large- r behavior of the wavefunctions, band energies versus total energies, volume integrated quantities versus explicit coordinate or momentum locally dependent quantities, etc.), we discuss here a multitude of properties. The reader is then to focus on aspects relevant to his interest. The detailed consideration of the accuracy of orbital energies, total energies, orbital moments, electrostatic potentials and energies and charge accumulation functions, allows considerable insight into the performance of these potentials in polyatomic systems. As a matter of fact, many of the shortcomings of these potentials in describing condensed systems can be identified and estimated in the atomic (or ionic) limit (see below). This is related to the fact that *relative* atomic energies are frequently characteristic of band energy *differences* in the solid, that a superposition of overlapping atomic charge densities

forms a good first order approximation to the crystalline density, and that nearly localized states in molecules and surfaces carry many of the characteristics of the atoms. To the extent that the atomic and ionic pseudo-wavefunctions mimic closely the "true" valence orbitals outside the core region, for a wide range of ground and excited states (i.e., configurations which can effectively mix in forming the solid), it is likely that bonding characteristics in polyatomic systems would be well reproduced. At the same time, atomic tests are unable to assess the reliability of the description of direct bonding effects in polyatomic systems (formation of bonding-antibonding pairs or directed bonds). This study hence complements similar comparisons for the Si_2 molecule¹⁻³ and bulk Si.⁴⁻⁵

We consider here the pseudopotentials that were recently applied to Si-based systems: the local semiempirical Berkeley potential (soft-core), SEMP¹² (which is nearly identical to the Appelbaum-Hamann potential¹³); the local model potential of Harris and Jones, MODEL;² the density functional potential, DF;²² and the trans-density functional potential, TDF.²⁷ They are displayed graphically in Fig. 1. All calculations are performed in the local density formalism with the standard Kohn and Sham exchange $\alpha = 2/3$. (Note that different choices of the exchange-correlation potential can produce significant changes in the results. However, the *error* of a given potential relative to the all-electron results is much smaller.)

A. Orbital energies and total energy differences

One would expect that a useful pseudopotential will reproduce the atomic orbital energies over a certain range of electronic configurations; at least those configurations that interact effectively in forming a solid. If one denotes by ΔE_i the difference in total energy of a certain configuration relative to the ground state, it is conceivable that the solid would be formed by mixing configurations that span a ΔE_i range characteristic of the scattering power of electrons in the solid. The latter is of the order of the band width W (or more).

Table I shows the Si orbital energies ϵ_{nl} and the total energy differences (relative to the $3s^23p^2$ configuration) obtained with the four different pseudopotentials, as well as the result obtained with the full all-electron (AE) approach. The errors in eV, relative to AE, are given in brackets. We include the $3s^13p^{2.5}3d^{0.5}$ configuration that has been used as a reference state to obtain the DF and TDF potentials (as bound $3d$ orbitals are needed to obtain the $l = 2$ potential). Figure 2 displays the errors in s and p energies as a function of the difference in total energy of a particular configuration from the ground atomic state, in units of the bulk Si band width W . We have included in the figure a number of intermediate excited states formed by changing the fractional occupancy of the orbitals, to obtain a more continuous description of the energy dependence.

It is seen that the SEMP potential places the $3s$ orbital about 0.6 eV too low relative to the $3p$ orbital at small $\Delta E_i/W$; at higher excitation states the description of the $3s$ becomes somewhat better whereas the $3p$ is too high. The *relative* errors remain of the order of 0.8 eV for intermediate and high excitations. Using an exchange coefficient of $\alpha = 1$ (for which the SEMP potential has been initially designed), the relative

TABLE I. Comparison of the Si orbital energies ϵ_{3s} , ϵ_{3p} , and ϵ_{3d} and the total energy differences ΔE_t (in eV), as obtained by the full all-electron (AE) approach and by various pseudopotentials. Numbers in brackets indicate the difference (in eV) relative to the AE results. TDF—nonlocal trans density functional; DF—nonlocal density functional; MODEL—the model local potential of Harris and Jones;² SEMP—the local semiempirical Berkeley potential.¹² 1 a.u. = 27.21 eV. All results are obtained with an $\alpha = 2/3$ coefficient.

	ϵ_{3s}	ϵ_{3p}	ϵ_{3d}	ΔE_t
$3s^2 3p^2$				
AE	-9.772	-3.206	...	0.0
TDF	-9.782 [0.01]	-3.209 [0.003]	...	0.0
DF	-9.702 [-0.07]	-3.161 [-0.04]	...	0.0
MODEL	-9.585 [-0.19]	-3.238 [0.03]	...	0.0
SEMP	-10.400 [0.63]	-3.289 [0.08]	...	0.0
$3s^1 3p^3$				
AE	-10.522	-3.764	...	6.664
TDF	-10.525 [0.003]	-3.766 [0.002]	...	6.669 [-0.005]
DF	-10.480 [-0.04]	-3.730 [-0.03]	...	6.645 [0.02]
MODEL	-10.305 [-0.22]	-3.782 [0.02]	...	6.437 [0.23]
SEMP	-11.307 [0.79]	-3.912 [0.15]	...	7.257 [-0.59]
$3s^1 3p^2 3d^{0.5}$				
AE	-13.279	-6.255	-0.199	9.184
TDF	-13.279 [0.00]	-6.255 [0.00]	-0.199 [0.00]	9.188 [-0.004]
DF	-13.279 [0.00]	-6.255 [0.00]	-0.199 [0.00]	9.158 [0.03]
MODEL	-12.977 [-0.30]	-6.195 [-0.06]	-0.199 [0.00]	8.945 [0.24]
SEMP	-14.032 [0.75]	-6.333 [0.08]	-0.199 [0.00]	9.835 [-0.65]
$\text{Si}^+ 13s^1 3p^2$				
AE	-18.571	-11.249	-3.249	14.038
TDF	-18.562 [-0.01]	-11.241 [0.008]	-3.250 [0.001]	14.041 [-0.003]
DF	-18.645 [0.07]	-11.311 [-0.06]	-3.251 [0.002]	14.027 [0.01]
MODEL	-18.111 [-0.46]	-11.044 [-0.21]	-3.034 [-0.22]	13.733 [0.305]
SEMP	-19.254 [0.68]	-11.206 [0.14]	-3.152 [-0.10]	14.027 [0.01]
$\text{Si}^{+2} 3s^1 3p^1$				
AE	-27.997	-20.086	-10.336	29.610
TDF	-27.932 [-0.06]	-20.032 [-0.05]	-10.320 [0.02]	29.589 [0.02]
DF	-28.307 [0.31]	-20.355 [0.27]	-10.389 [0.05]	29.757 [-0.15]
MODEL	-27.077 [-0.92]	-19.446 [-0.64]	-9.390 [-0.95]	28.904 [0.71]
SEMP	-28.307 [0.31]	-19.525 [-0.56]	-9.607 [-0.73]	-30.008 [-1.40]

errors are of the order of 1 eV or more.^{5,20} This relative error is essentially carried over to the solid and the Si₂ molecule. The error in the *s*–*p* spacing with SEMP is 0.8 eV in the solid^{4,5} and about 0.7 eV in the molecule.³

The model potential has smaller errors: of the order of $-(0.2\text{--}0.5)$ eV for *s* and $-(0.02\text{--}0.2)$ eV for *p*. These errors have, however, the same sign over a wide energy range and are expected and found⁴ to partially cancel in the band

structure of the solid. Note, however, that this cancellation comes to effect only because the *absolute* energies with respect to vacuum are usually of no interest in describing the band structure. In the calculation of the work function or the total crystal energy, however, no such cancellation is expected to occur. The errors associated with the DF and the TDF potentials are below 0.1 eV for $\Delta E_t/W \leq 1.5$.

B. Orbital moments

A direct graphical display of the wavefunctions is not effective for gauging the relevant errors, as volume elements are not easily perceived. Instead, we use a number of integrated quantities. The first of them is the *p*th orbital moment $\langle \chi_{nl} | r^p | \chi_{nl} \rangle$. These form a quantitative measure to the orbital localization and determine the electrostatic potential set in the system by the density $\chi_{nl}^2(r)$.

Table II shows the orbital moments for *p* = 1 and 2. Percent errors relative to the AE results are given in brackets.

The SEMP orbitals are seen to be too spread out: the first moment is too large by about 4% to 9%, while the second moment (emphasizing the tail behavior) has even larger errors. This is also borne out by the calculated bulk charge density^{4,5} that indicates a concentration of charge in the bond center at the expense of the intermediate-*r* regions. The model potential has larger errors in the same direction. The errors in the DF moments are smaller than both the SEMP and the

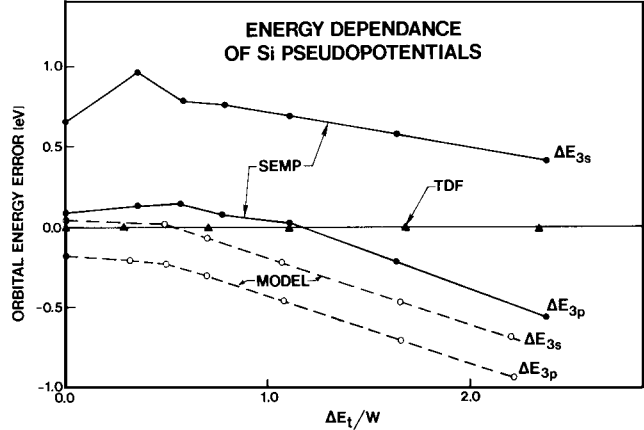


FIG. 2. The errors in the pseudopotential orbital energies, ϵ_{3s} and ϵ_{3p} , relative to the all-electron values, displayed as a function of the excitation energy of each configuration, $\Delta E_t/W$ (in units of the bulk Si valence band width *W*).

TABLE II. Comparison of the s and p pseudo-orbital moments with the all-electron moments. Numbers in brackets indicate percent error relative to the all-electron results. Notation as in Table I.

	$\langle r \rangle_{3s}$ (au)	$\langle r \rangle_{3p}$ (au)	$\langle r^2 \rangle_{3s}$ (au ²)	$\langle r^2 \rangle_{3p}$ (au ²)
$3s^2 3p^2$				
AE	2.178	2.877	5.555	10.083
TDF	2.180 [−0.1%]	2.887 [−0.35%]	5.553 [0.04%]	10.081 [0.02%]
DF	2.110 [3.1%]	2.840 [1.3%]	5.337 [3.9%]	9.839 [2.4%]
MODEL	2.300 [−5.6%]	2.993 [−4.0%]	6.031 [−8.6%]	10.665 [−5.7%]
SEMP	2.311 [−6.1%]	3.009 [−4.6%]	6.738 [−3.3%]	10.712 [−6.2%]
$3s^1 3p^3$				
AE	2.144	2.768	5.369	9.268
TDF	2.157 [0.6%]	2.780 [−0.4%]	5.373 [0.07%]	9.278 [−0.1%]
DF	2.073 [3.3%]	2.717 [1.8%]	5.137 [4.3%]	8.968 [3.2%]
MODEL	2.270 [−5.9%]	2.894 [−4.6%]	5.858 [−9.1%]	9.887 [−6.7%]
SEMP	2.230 [−4.0%]	2.897 [−4.7%]	5.614 [−4.6%]	9.839 [−6.2%]
$3s^1 3p^2$				
AE	2.052	2.450	4.866	7.054
TDF	2.067 [−0.7%]	2.467 [−0.7%]	4.875 [−0.2%]	7.077 [−0.3%]
DF	1.976 [3.7%]	2.387 [2.6%]	4.624 [5.0%]	6.701 [5.0%]
MODEL	2.185 [−6.5%]	2.598 [−6.0%]	5.359 [−10.1%]	7.722 [−9.5%]
SEMP	2.203 [−7.3%]	2.620 [−7.0%]	5.213 [−7.1%]	7.817 [−10.8%]
$3s^1 3p^1$				
AE	1.954	2.232	4.380	5.782
TDF	1.973 [−1.0%]	2.246 [−0.6%]	4.400 [−0.4%]	5.825 [−0.7%]
DF	1.874 [4.1%]	2.163 [3.1%]	4.127 [5.8%]	5.412 [6.4%]
MODEL	2.099 [−7.4%]	2.405 [−7.7%]	4.902 [−12%]	6.503 [−12.5%]
SEMP	2.134 [−9.6%]	2.437 [−9.2%]	4.532 [−3.5%]	6.667 [−15.3%]

MODEL errors, but are in the opposite direction: the orbitals are too localized. This reflects the fact that the unitary rotation in the space of the AE orbitals [e.g., Eq. (9)] introduces core components (i.e., $C_{n,n'}^{(l)} < 1$) with their characteristically smaller orbital moments. This deficiency is directly addressed in the TDF potential, which leads to errors below 1%.

Note that as sometimes small errors in the orbital energies are associated with fairly large errors in the wavefunctions, both aspects need to be examined. The negative moment ($p = -1$) obtained with all pseudopotentials contains very large errors (10%–20%) indicating the inadequacy of the pseudo-potential approach for describing the core region.

C. X-ray scattering factors

Another quantity that measures the quality of the pseudo charge density is its Fourier transform. Figure 3 displays the error,

$$\Delta F(q) = \int e^{iqr} \rho_v(r) dr - \int e^{iqr} n(r) dr \quad (15)$$

associated with the pseudo x-ray scattering factor, relative to the valence contribution of the corresponding all-electron quantity. The $3s^1 3p^3$ configuration is used. The momentum q is a continuous variable for the atom. The vertical bars in Fig. 4 indicate the experimental errors in determining the bulk Si scattering factors at the few lowest reciprocal lattice vectors. Both the SEMF and the MODEL potential have errors that are larger than the experimental uncertainty in the physically important low-momentum region. The sign of the error indicates again that the charge is too spread out in the bond region. The DF potential has lower errors, but nevertheless nonnegligible. (Note that upon core-orthogonalization, these errors vanish identically for the DF case.^{5,22}) In the TDF approach, the errors are reduced substantially and are pushed

to the high-momentum regions that are less important in describing the valence properties of condensed systems.

D. Charge accumulation functions

As the various pseudopotentials discussed above have widely different forms in the core region, they lead to different values for the amount of charge enclosed in the core and consequently to different valence charges. These effects are conveniently monitored by considering the orbital charge accumulation function:

$$Q_{nl}(\mathbf{R}) = \int_0^{\mathbf{R}} \chi_{nl}^*(\mathbf{r}) \chi_{nl}(\mathbf{r}) d\mathbf{r} \quad (16)$$

$$Q_{nl}(\infty) = 1.0.$$

Note that $Q_{nl}(\mathbf{R})/\mathbf{R}$ determines the electrostatic potential. Figure 4 displays $Q_{3s}(\mathbf{R})$ and $Q_{3p}(\mathbf{R})$ for Si, calculated with the various pseudopotentials for the ground state configuration $3s^2 3p^2$. Figure 5 displays the percent errors for another configuration ($3s^1 3p^3$). The positions of the bond center in the solid and the outer atomic nodes are indicated by arrows.

It is seen that the model potential accumulates charge much more slowly than the all-electron case and that errors of 10% to 20% persist between the outer nodes and the bond center. On the other hand, the DF potential accumulates charge too fast and shows errors of $-(5-10)\%$ in the same region. The TDF potential yields very small errors of $-(2-10^{-3})\%$: its charge accumulation functions join the all-electron results at $1.2a_0$.

E. Coulomb integrals and the total energy problem

We have so far considered quantities that reflect the accuracy of the orbital energies, electrostatic potentials, and

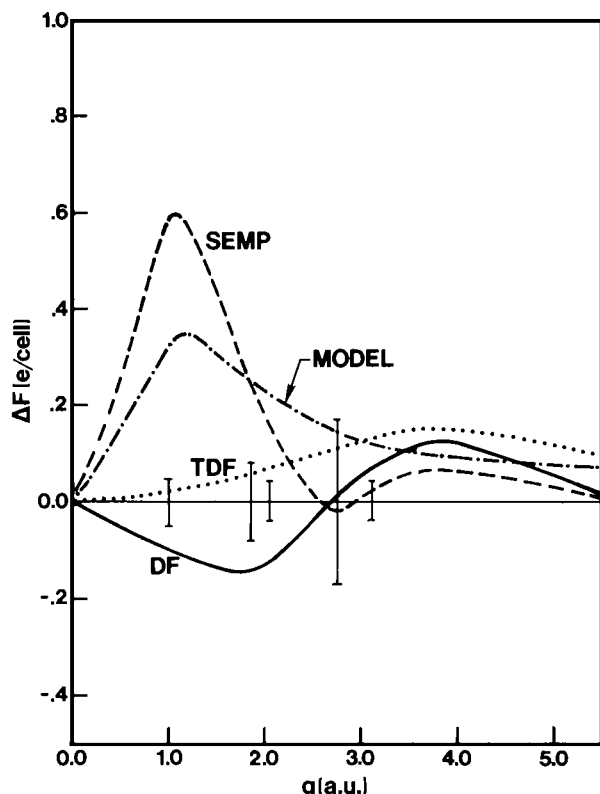


FIG. 3. The errors in the pseudopotential x-ray scattering factor, relative to the all-electron results. The vertical bars indicate the magnitude of the experimental uncertainty in determining the respective quantities in the bulk, at the first few reciprocal lattice vectors.

charge densities. If one is interested directly in the total energies, the convolution of the latter two quantities becomes important. This is the Coulomb integral:

$$J_{nl,nl} = \int \frac{\rho_{nl}(\mathbf{r})\rho_{nl}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}d\mathbf{r}' \quad (17)$$

where $\rho_{nl}(\mathbf{r})$ is the orbital charge. Table II presents these in-

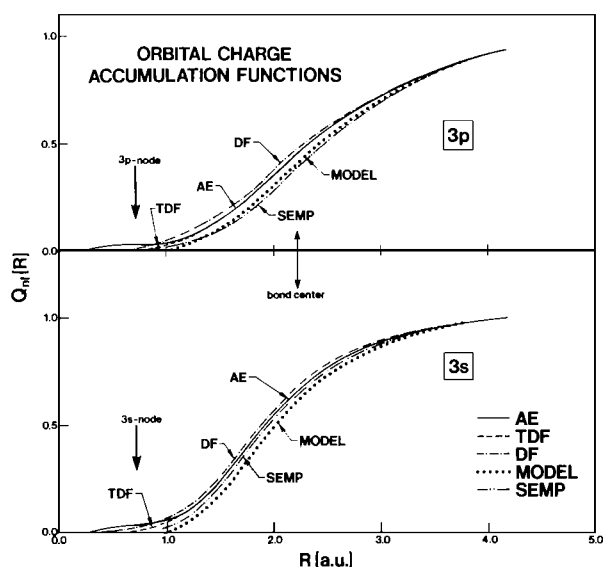


FIG. 4. Orbital charge accumulation functions for Si. $Q_{nl}(R) = \int_0^R \chi_{nl}(r) dr$. The vertical arrows indicate the position of the last radical node in the atom and the bond center in the bulk solid. The $3s^1 3p^{2.5} 3d^{0.5}$ configuration is used.

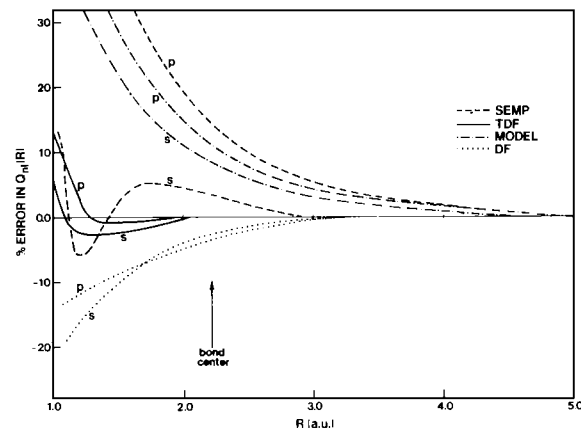


FIG. 5. Percent error in the orbital charge accumulation functions of Si, relative to the all-electron results, for the $3s^1 3p^3$ configuration. The vertical arrow indicates the position of the bond center in the bulk solid.

tegrals as calculated from the four pseudopotentials and the all-electron case.

It is seen that the SEMP and MODEL pseudopotentials produce values for $J_{nl,nl}$ that are considerably too large (in particular, a large error occurs in $J_{3d,3d}$, for which nonlocal corrections are important). The DF and TDF potentials produce errors smaller than 1%.

Note that in general it is not sufficient to ensure the correct value of the pseudopotential Coulomb (and exchange) integrals in order to get an accurate total energy, as in the pseudopotential approach, the core-attraction is also modified relative to the all-electron case. This leads to the possibility that the errors in the *core attraction* terms of "soft" pseudopotentials, introduced by the penetration of excess valence charge into the core (and reflected by overly attractive orbital energies, cf. Table I and similar errors in the band structure) will be partially offset by the errors in the *Coulomb repulsion* introduced by the accumulation of excess charge in the bond region (as reflected by the over-repulsive Coulomb energies, cf. Table III). Hence, as the stationary character of the total variational energy permits partial cancellation of errors which originate from different spatial regions in the density matrix, the final soft-core pseudopotential total energy may well be more accurate than the errors in its individual components. The DF and TDF pseudopotentials attempt, on the other hand, to reproduce accurately the *individual constructs* of the total energy expression as they constitute by themselves physically interesting observables (e.g., orbital energies, charge density).

F. Summary of results

We conclude from these tests the following:

(i) The SEMP "soft-core" potential^{12,13} produces overly diffused orbitals [e.g., the positive orbital moments are overestimated by 4% to 9%, while the low-momentum x-ray structure factors are significantly underestimated, (Fig. 3)], yielding differences as large as 15% in the charge accumulation functions at the bond region. This leads to a somewhat distorted bulk⁴ and molecular¹ charge density. The errors in the relative *s* and *p* orbital energies are of the order of 0.8 eV, also characteristic of the differences found in molecular and

TABLE III. Calculated Coulomb integrals (in Hartrees). Notation as in Table I. Percent errors relative to the AE results are given in brackets.

	$J_{3s,3s}$	$J_{3p,3p}$	$J_{3d,3d}$
AE	0.4299	0.3490	0.0560
TDF	0.4291 [−0.2%]	0.3480 [−0.3%]	0.0560 [0.0%]
DF	0.4342 [−1.0%]	0.3482 [−0.2%]	0.0560 [0.0%]
MODEL	0.4325 [−0.6%]	0.3567 [−2.2%]	0.0613 [−9.5%]
SEMP	0.4401 [−2.4%]	0.3665 [−5.0%]	0.0622 [−11.1%]

solid state calculations with the same potential. The energy dependence of these errors remain large over a wide excitation range. The over-attractiveness of the electron-core potential is partially offset by the over-repulsiveness of the Coulomb energy, suggesting that the error in the total energy might be considerably smaller than these individual errors.

(ii) The local model potential² produces smaller errors in the orbital energies (≤ 0.5 eV) over an excitation range of about 1 Ryd. These errors have the same sign for s and p and tend to cancel in the conventional band structure representation. The associated orbitals are too diffused (e.g., they overestimate the positive orbital moments by 5% to 12% and underestimate the charge accumulation functions by 15% to 20% in the bond region). This potential is accurate near the ground state configuration, while for excited states with d -character (e.g., Γ_{12} in bulk Si) errors of 1 eV are apparent.⁴

(iii) The nonlocal DF pseudopotential²² yields yet smaller errors (≤ 0.1 eV) in the orbital energies for an excitation range of more than 1 Ryd. The errors are usually of the same sign and have small energy dependences. The associated orbitals are too localized (e.g., they underestimate the positive orbital moments by 2% to 5% and overestimate the charge accumulation functions by 10%).

(iv) The trans-density-functional nonlocal potentials have errors smaller than 0.01 eV in the atomic orbital energies over a 1-Ryd-wide excitation energy range and reproduce the orbital moments, charge accumulation functions, and Coulomb energies to within 1% over the entire range. They should prove superior to the DF potential when very high accuracy is needed (e.g., the calculation of molecular binding curves). For other purposes, the DF pseudopotential is sufficiently accurate. Recent tests on bulk Si^{5a} indicate that the TDF potential reproduces very well the AE valence charge density and yields a band structure with an average error of only 0.06 eV/state, over a 20 eV range, relative to the full calculation.

IV. SOME COMMON MISCONCEPTIONS REGARDING PSEUDOPOTENTIALS

We close this paper by commenting on a number of questions, some more trivial than others, that have apparently introduced some difficulties in discussing the validity of the pseudopotential description of semiconductors and surfaces. This section is not intended to provide a thorough theoretical discussion of the problem but rather attempts to elucidate in a *simple* manner some of these problems for readers who are not entirely familiar with the theoretical mumbo-jumbo in the field.

(1) *Pseudopotentials are “weak”*: Pseudopotentials are “weak” only in one sense: the orbital energies of their lowest states (i.e., valence) are higher than that of the corresponding all-electron problem (having core states). They are not necessarily “weak” in their spatial behavior; as a matter of fact some successful pseudopotentials are even singular! There are certain clear regularities in the degree of repulsiveness of atomic potentials as one moves along rows and columns in the periodic table. These changes reflect genuine chemical effects and are useful in describing many properties of condensed phases, such as the prediction of the stable crystal structure of (about 460) AB Compounds.²⁶ Pseudopotentials *can be made weak* (or soft) if one desires but in general there exists a continuous tradeoff between their weakness and the accuracy of the associated wavefunctions.

(2) *A good atomic pseudopotential should attempt to closely describe the “valence” region while its form in the core region is irrelevant*: This notion is rooted in the concept that since in the pseudopotential representation the core region is being “*pseudized away*” and the true nodal character of the wavefunction in this region is eliminated, one can use any convenient form for both the wavefunction and the pseudopotential in the core. As indeed there is no physically compelling *simple* rule dictating the shape of the pseudopotential in this region (nor is there a sum rule to be satisfied at the high-momentum short-wavelength limit), one is faced with a tremendous abundance of pseudopotentials in the literature, differing predominantly in their shape in the core region (square wells, Gaussian-shaped, soft-core, hard-core, empty-core, etc.) These potentials are frequently judged only by their closeness to the “true” potential in the valence region.

The form of the pseudopotential in the core region has, however, an enormous bearing on its *transferability* from the model system used for its construction (e.g., an atom or bare ion) to the systems of interest (molecules, solids and surfaces). A pseudopotential is, of course, useful only to the extent that it can continue to approximate well the full all-electron results for atomic sites having symmetries and chemical environments that are sometimes very different from that of the systems used to “fit” or derive it. The shape of the pseudopotential in the core region largely determines this energy dependence. Hence, for example, soft pseudopotentials are associated with wavefunctions having fairly large amplitudes in the core regions and are loosely tied at the origin. Although one can trivially construct such a potential to yield an acceptable description of the reference system, when the latter is placed in a different environment, there is no mechanism in an unconstrained self-consistent solution of the pseudo-Schrödinger equation that will prevent the *valence* charge from overpenetrating the soft-core region, destroying thereby the agreement with the correct valence implied by the fitting procedure. A suitably constructed hard-core pseudopotential, on the other hand, first imposes an accurate description of the valence region and then “ties down” the pseudo-wavefunctions at the origin to ensure a minimal core penetration under chemical perturbations. Scattering events at lower energies, which reflect the changed chemical environment (hybridization, screening, formation of directed and localized bonds,

etc.) are then described by the valence regions of the wavefunctions.

Figure 2 shows a simple demonstration of these effects in an atomic case: as the valence structure of the atom is changed from its reference ground state ($\Delta E_l = 0$) through excitations, soft potentials continue to reflect sizable errors in the orbital energies, while hard potentials deteriorate much more slowly. A recent calculation on Si_2 with soft and hard potentials^{1,3} demonstrate a similar effect: when a soft potential, adjusted to fit some quantities in a (spherical) Si ion, is used to calculate the electronic structure of a (cylindrical) Si_2 molecule, a considerable amount of charge penetrates the molecular core, thereby partially offsetting the agreement with the all-electron results in the *bond* region. An *ad hoc* core orthogonalization³ (that acts as an added repulsive potential) is then needed to restore the correct topology of the charge in the bond region (effects that are sometimes as high as 15% to 20%).

(3) *“Soft” pseudopotentials are more convenient to use than “hard” pseudopotentials:* This notion is rooted in the practice of expanding the system's wavefunction in plane waves: a “soft” potential converges, then, faster in momentum space and, hence, fewer matrix elements and smaller secular equations need to be considered.

Although the plane-wave-based band structure technology has been a long-time favorite, it is by no means advantageous over (nonperturbational) methods using direct-space basis functions (analytic Gaussians or Slaters, numerical basis functions, radial partial waves, etc.). Using the latter, hard-core potentials are treated as easily as soft-core potentials (e.g., Ref. 16–21). Moreover, systems for which a plane wave expansion converges unacceptably slowly (e.g., systems containing first row atoms or 3d transition metals) can be treated conveniently by direct-space-based basis functions. There is therefore no compelling computational reason at present to construct soft-core potentials for self-consistent electronic structure calculations.

Soft-core pseudopotentials have enjoyed an enormous success in conjunction with low order perturbation theory approaches to transport and structural properties of solids.⁷ As the conventional choice of the unperturbed reference system is a homogeneous electron gas, an adequate convergence of the perturbation series as well as that of the relevant reciprocal lattice sums rests heavily on the suitability of a soft-core pseudopotential representation to the description of the electron-core scattering events. Whereas this situation is approximately realized in simple metals (notable exceptions include calculated quantities which involve nonzero order *wavefunctions*, e.g., optical conductivity, etc.), many of the chemically interesting covalent as well as transition metal systems are not amenable to such a simplified description. It would seem reasonable to expect that hard-core potentials could be used successfully in such situations provided the nonperturbed reference system is chosen as suitably *nonhomogeneous* (e.g., noninteracting pseudo-atoms, square wells, etc.), and hence incorporate most of the potential fluctuations at the zero-order level.

(4) *The misrepresentation of the Si bulk charge density by a local (semiempirical) pseudopotential results from lack*

of nonlocality effects: Early⁸ self-consistent calculations with the soft semiempirical local Si pseudopotentials (see also Ref. 5) have indicated that the bulk charge density has certain anomalies relative to the experimentally synthesized density²⁹: the calculated charge, instead of being prolated along the Si–Si bond, is oblate and looks as if it points perpendicular to the bond. A subsequent calculation by Chelikowsky and Cohen,¹⁰ this time using nonself-consistent empirical pseudopotentials, have shown that the addition of a nonlocal term represented as a two parameter Gaussian restores the expected shape of the density and yields at the same time a band structure in excellent agreement with experiment. Although never stated by these authors, it has often been suggested in the literature that some kind of nonlocality is necessary for a correct description of the Si covalent charge density. This, however, is not necessarily the case.

All atomic pseudopotentials that can adequately replace an all-electron problem are inherently nonlocal. This nonlocality can (but does not have to) be conveniently represented as an angular momentum dependence: $V_{ps}^{(l)}(\mathbf{r})$; $l = 0, 1, 2$. The validity of the replacement of $V_{ps}^{(l)}(\mathbf{r})$ by a local form acting equally on all the electrons depends on the shell structure of the atom (that is partially reflected by its position in the periodic table). First-principles calculations of the full nonlocal pseudopotentials of all atoms belonging to the first five rows in the periodic table²² have offered quantitative measures to the importance of the nonlocality in various parts of the periodic table. They also show that nonlocal effects for atoms such as Si, Ga, As, Ge, etc. are not nearly as enormous as they are for first row atoms or the 3d transition metal series, and hence suitable local forms may be adequate.

It has been demonstrated^{3,4} that one can obtain the approximately correct (molecular or bulk) Si bond charge anisotropy even with local potentials, provided these are constructed by requiring a good reproduction of both the all-electron orbital energies *and the wavefunctions* in the valence region. This more general approach results, however, in “harder” and more localized pseudopotentials than previously used (e.g., the average of the first principles *s* and *p* potentials in Fig. 1). Hence, although one can obtain the expected bond charge anisotropy by additional fitting of an otherwise local potential,¹⁰ the physical origin of this anisotropy seems to be related to the small-*r* localization of the potential that in turn is a consequence of constraining it to closely reproduce the all-electron charge density. We note that while suitable local potentials can be adequate for such systems, an accurate description of the electronic structure (in particular *s*–*p* hybridization effects, role of *d*-states, anisotropy of dangling bonds, etc.) requires the full non-local potential.

(5) *Core-orthogonalized pseudo wavefunctions give a better description of the electronic structure than pseudo-wavefunctions:* This is true only if the pseudo wavefunction can be represented (at least approximately) as a linear combination of the all-electron core plus valence wavefunctions [e.g., Eq. (9)]. Core projections then act to annihilate the core part, yielding the true (nodal) valence wavefunctions that can be used to obtain the internally consistent theoretical predictions of various physical observables. As the use of pseudopotentials with their associated smooth wavefunctions

constitutes an enormous simplification over the direct solution of the all-electron problem, such a core orthogonalization (performed *after* the completion of the solution of the pseudopotential problem) offers great simplicity and very reasonable accuracy.

Note, however, that in the empirical and semiempirical pseudopotential approaches one does not have a direct handle on the description of the pseudo-orbitals in terms of the all-electron orbitals (as one parametrizes the pseudopotential to yield desired energy eigenvalues without specifying the space spanned by the corresponding orbitals) and, hence, a core orthogonalization is at best an *ad hoc* procedure. As a matter of fact, in some cases (e.g., form factors of the Si atom^{5,22}) the core orthogonalized orbitals derived from a semiempirical potential yield poorer results than the unorthogonalized orbitals. First principles pseudopotentials having orbitals that are directly constrained to span the core and valence space underlying the all-electron system can be used in conjunction with the orthogonalization procedure. These yield consistently better results upon core orthogonalization.²²

If one desires to avoid altogether the need for core orthogonalization and still obtain approximately valid results at this level, one needs to construct a pseudopotential that yields orbitals with a minimum amplitude in the core region. This is precisely what is achieved with “hard-core” pseudopotentials^{2,22} which explicitly minimize the core projections (i.e., the core orthogonality terms).

(6) *Nonlocal pseudopotentials can represent the electronic structure of a system better than an all-electron Kohn and Sham approach as the latter is a local approximation:* This notion argues that as the all-electron local density functional formalism employs a single potential for all the states in the system, a nonlocal pseudopotential is “better.” We note, however, that even a local *all-electron* potential is equivalent to an infinite series of *l*-dependent (nonlocal) *pseudopotentials*. Hence, two different types of nonlocalities are being confused: that inherent to the description of the interelectronic exchange and correlation effects and manifested already at the all-electron level by terms such as the nonlocal Hartree–Fock exchange, gradient contributions in the density functional approach, etc. (“essential nonlocality”); and the pseudopotential nonlocality that is a simple mathematical consequence of eliminating a certain orthogonal subspace from the eigenvector spectrum. *An internally consistent nonlocal pseudopotential can hence be at best as good as the underlying all-electron approach.*

There is a large number of pseudopotential band structure results in the literature that were shown to disagree with all-electron results, even though both approaches use the same microscopic description of screening effects (e.g., $X\alpha$). These differences often form the basis for the discussion on the superiority of one approach over the other. Indeed, empirical or semiempirical pseudopotentials are derived from sources that are often inconsistent with the approach used to construct their corresponding screening (*viz.*, empirically adjusted ionic potentials screened by $X\alpha$ potentials) and, hence, should not be contrasted with all-electron results that employ consistently the same screening for the core as well as the valence electrons. First-principles pseudopotentials, on the other hand, are di-

TABLE IV. First principals density-functional results for the band gaps of some cubic materials compared with experimental data.

Material	Calculated gap (eV)	Experimental gap (eV)	Percentage difference
Si	0.55 ^{a,b}	1.1	50
LiF	10.5 ^c , 10.4 ^d	14.2	26
CuCl	2.0 ^c	3.2	37
CdS	2.0 ^f	2.55	22
GaAs	1.0 ^g	1.4	28

^a All-electron, linear augmented plane wave method, Ref. 4.

^b First principles pseudopotential, Ref. 5a.

^c Gaussian basis LCAO, Ref. 32.

^d Numerical LCAO, Ref. 33.

^e Numerical LCAO, Ref. 34.

^f Numerical LCAO, Ref. 35.

^g First principles pseudopotentials, Ref. 25.

rectly derived from a chosen screening theory (Hartree–Fock, density functional, etc.) and have to be judged by the degree they reproduce the corresponding all-electron results.

(7) *There is no direct way of abstracting the total energy of a system using semiempirically adjusted pseudopotentials:* The prediction of crystal structure, relaxation and surface reconstruction effects involves the minimization of the total systems energy with respect to structural coordinates. This can be effectively done in cluster-type Hartree–Fock or correlated Hartree–Fock pseudopotential approaches.²¹ The same can be achieved using semiempirical pseudopotentials.³⁰ A suitable and convenient formalism to do that for extended systems in a momentum space representation has been recently developed³¹ and applied successfully to bulk and surface Si.³⁰ Note, however, that this leads to some theoretical inconsistencies in describing correlation effects: the fitting of the potential to yield the experimentally observed ionic term values introduces some unspecified form of correlation that is absent from the screening energy introduced via a certain (imperfect) correlation model (e.g., Kohn and Sham). The use of a pseudopotential that is internally consistent with a corresponding microscopic screening model (density functional,²² Hartree–Fock,^{18–19} etc.) is free from this inconsistency and yields reasonable results.^{17,18,24}

(8) *State-of-the-art first-principles band structures (pseudopotential or all-electron) yield excellent agreement with experimental optical data:* Table IV compares the predicted band gaps^{32–35} of a number of solids, obtained with state-of-the-art non-empirical band structure techniques, and with the observed data. The calculations represented in the table are fully self-consistent, avoid any shape approximation to the potential, and use highly converged representations for the wavefunction and potential. Their internal accuracy is said to be of the order of 0.1–0.2 eV or better. On this scale, the discrepancies with experiment are enormous. There is no reason to believe that these errors will be reduced for surface calculations. This signals the breakdown of the one-electron Koopmans-type band structure approach for a wide class of materials.

The physical origin of some of these discrepancies and the evaluation of the relevant many-body correction to them are discussed elsewhere (e.g., Ref. 33). These corrections are by no means easy to evaluate and require an approach that fre-

quently abandons the simplicity of the band structure model. It would seem that one of the more important challenges of the contemporary many-body theory is to construct (seemingly) one-body potentials that would yield consistently better predictions for the elementary excitations, already at the simple eigenvalue equation level, without sacrificing the accuracy of the predicted ground state properties. Empirical and semiempirical pseudopotential approaches avoid this complexity by a phenomenological adjustment of parameters. The insight offered by them into the nature of the microscopic interactions in the system remains limited.

ACKNOWLEDGMENTS

I am pleased to thank L. R. Kahn for many helpful discussions and to D. R. Hamann for calculating the Si bulk band structure with the DF potential and discussing the results.

- ¹D. J. Miller, D. Haneman, E. J. Baerends, and P. Ros, *Phys. Rev. Lett.* **41**, 197 (1978).
- ²J. Harris and R. O. Jones, *Phys. Rev. Lett.* **41**, 191 (1978).
- ³M. Schlüter, A. Zunger, G. P. Kerker, K. M. Ho, and M. L. Cohen, *Phys. Rev. Lett.* **42**, 540 (1979).
- ⁴D. R. Hamann, *Phys. Rev. Lett.* **42**, 662 (1979).
- ⁵(a) D. R. Hamann and A. Zunger (unpublished). (b) A. Zunger and M. L. Cohen, *Phys. Rev. B* **18**, 5449 (1978).
- ⁶D. J. Miller and D. Haneman, *Phys. Rev. B* **17**, 5033 (1978); *J. Vac. Sci. Technol.* **15**, 1267 (1978).
- ⁷E. G. L. Kleinman and J. C. Phillips, *Phys. Rev.* **118**, 1153 (1960); W. A. Harrison in *Pseudopotential In the Theory of Metals* (Benjamin, New York, 1966).
- ⁸M. L. Cohen and V. Heine, *Solid State Phys.* **24**, 38 (1970).
- ⁹D. Brust, *Methods Comput. Phys.* **8**, 33 (1968).
- ¹⁰J. R. Chelikowsky and M. L. Cohen, *Phys. Rev. B* **10**, 5095 (1974).
- ¹¹J. R. Chelikowsky and M. L. Cohen, *Phys. Rev. B* **14**, 556 (1976).
- ¹²M. L. Cohen, M. Schlüter, J. R. Chelikowsky, and S. G. Louie, *Phys. Rev. B* **12**, 5575 (1975).
- ¹³J. A. Appelbaum and D. R. Hamann, *Phys. Rev. B* **8**, 1777 (1973).
- ¹⁴Rewriting the pseudopotential single-particle Eq. (2) in the form:

$$W_{ps}^{(j)}(\mathbf{r}) = \epsilon_j + \frac{1}{2} \frac{\nabla^2 \chi_j(\mathbf{r})}{\chi_j(\mathbf{r})} - V_o \left[\sum_j \chi_j^2(\mathbf{r}) \right]$$
 it is clear that the adjustment of a chosen parametrized (local) form of $W_{ps}(\mathbf{r})$ to reproduce a certain set of ϵ_j s, can be achieved with widely different wavefunctions χ_j . Without further sufficient constraints on χ_j , the imposition of an arbitrary form for $W_{ps}(\mathbf{r})$ might hence result in unwarranted forms for the wavefunctions.
- ¹⁵J. A. Appelbaum and D. R. Hamann, *Rev. Mod. Phys.* **48**, 3 (1976).
- ¹⁶S. G. Louie, *Phys. Rev. Lett.* **40**, 1525 (1978).
- ¹⁷W. A. Goddard, *Phys. Rev.* **174**, 659 (1968).
- ¹⁸L. R. Kahn, P. Baybutt and D. G. Truhelar, *J. Chem. Phys.* **65**, 3876 (1976).
- ¹⁹L. R. Kahn and W. A. Goddard, *Chem. Phys. Lett.* **7**, 667 (1968).
- ²⁰P. J. Hay, W. R. Wadt and L. R. Kahn, *J. Chem. Phys.* **68**, 3059 (1978).
- ²¹A. Redondo, W. A. Goddard and T. C. McGill, *Phys. Rev. B* **15**, 5038 (1977); W. A. Goddard, J. J. Barton, A. Redondo, and T. C. McGill, *J. Vac. Sci. Technol.* **15**, 1274 (1978).
- ²²A. Zunger and M. A. Ratner, *Chem. Phys.* **30**, 413 (1978); S. Topiol, A. Zunger, and M. A. Ratner, *Chem. Phys. Lett.* **49**, 367 (1977).
- ²³G. P. Kerker, A. Zunger, M. L. Cohen, and M. Schlüter (unpublished).
- ²⁴A. Zunger and M. L. Cohen, *Phys. Rev. B* (in press); A. Zunger, G. P. Kerker, and M. L. Cohen, *Phys. Rev. B* (in press).
- ²⁵A. Zunger (unpublished).
- ²⁶A. Zunger and M. L. Cohen, *Phys. Rev. Lett.* **41**, 53 (1978), A. Zunger (unpublished).
- ²⁷A. Zunger and L. R. Kahn (unpublished).
- ²⁸D. R. Hamann, private communication to A. Zunger, for which I am grateful.
- ²⁹Y. W. Yang and P. Coppens, *Solid State Commun.* **15**, 1555 (1974).
- ³⁰J. Ihm and M. L. Cohen (unpublished).
- ³¹J. Ihm, A. Zunger, and M. L. Cohen (unpublished).
- ³²W. P. Menzel, C. C. Lin, D. Fouquet, E. E. Lafon, and R. C. Chaney, *Phys. Rev. Lett.* **30**, 1313 (1973).
- ³³A. Zunger and A. J. Freeman, *Phys. Rev. B* **16**, 2901 (1977).
- ³⁴A. Zunger and M. L. Cohen, *Phys. Rev. B* (in press).
- ³⁵A. Zunger and A. J. Freeman, *Phys. Rev. B* **17**, 4850 (1978).