Nonlocal pseudopotential calculation of the electronic properties of relaxed GaAs (110) surface

Alex Zunger
Solar Energy Research Institute, Golden, Colorado 80401
(Received 26 December 1979)

The electronic structure of a semiconductor surface is studied for the first time using self-consistent nonlocal (first-principles) pseudopotentials. In agreement with the recent local pseudopotential as well as tight-binding studies, no intrinsic surface states are obtained in the gap of GaAs for the relaxed surface. However, in contrast with the previous approaches, new features of the electronic structure are obtained, including a pronounced downwards displacement of the low As-derived surface states, the appearance of an additional As-p state near the valence-band maximum, the reordering of the states near \( \Gamma \) with a different order of wave-function parity, and the development of pronounced d-orbital character (in addition to s and p) in the highest occupied and lowest empty surface states.

I. INTRODUCTION AND CONCLUSIONS

All of the pseudopotential electronic-structure calculations of semi-infinite semiconductor surfaces known today have used local pseudopotentials. On the other hand, all pseudopotentials are inherently nonlocal.\(^1\) Until recently, nonlocal pseudopotentials which are usable in self-consistent calculations were not available. Empirical nonlocal-like corrections to otherwise local pseudopotentials\(^2\)\(^-\)\(^3\) have shown substantial changes both in the band structure as well as in the electronic charge densities of zinc-blende semiconductors.\(^2\) Since, however, in this approach the screened pseudopotential, rather than the bare pseudopotential, has been empirically parametrized,\(^2\) these potentials were not expressible in terms of the calculated wave functions and hence permitted neither self-consistency in the calculation nor the transferability from bulk to surface calculations.

Recently,\(^4\) self-consistent first-principles nonlocal pseudopotentials have been derived from the density-functional formalism\(^5\) for all atoms of the first five rows in the periodic table. They have been used for a number of electronic-structure calculations, including that of bulk semiconductors,\(^6\) bulk transition metals,\(^6\)\(^-\)\(^7\) and diatomic molecules,\(^8\)\(^-\)\(^9\) as well as the calculation of the equilibrium lattice constant, cohesive energy, and bulk moduli of Si (Ref. 10) and Mo, W.\(^11\) In this paper, these nonlocal pseudopotentials are applied for the first time to study self-consistently the electronic structure of the relaxed GaAs (110) surface. I use the 27°-rotation relaxation model suggested by Tong et al.\(^12\) for the surface geometry.

My main conclusions are:

1. In agreement with recent tight-binding\(^13\)\(^-\)\(^14\) and local pseudopotential\(^15\) calculations using the relaxed\(^15\) surface geometry, but in contrast with the calculations using the ideal surface geometry,\(^16\)\(^-\)\(^19\) I find that no intrinsic surface states exist in the fundamental gap. This agrees with recent experimental findings.\(^20\)\(^-\)\(^25\) Earlier experiments on Fermi-level pinning\(^22\) as well as partial-yield spectroscopy\(^27\) have erroneously led to the conclusion that intrinsic surface states exist in the gap. However, subsequent studies on pure samples\(^28\)\(^-\)\(^29\) have indicated that no pinning exists in the gap, and that the core-to-empty-state transitions observed in partial-yield spectroscopy actually measured final core exciton states rather than one-electron surface states.\(^27\)\(^-\)\(^28\) The present study, based on nonlocal pseudopotentials, shows, however, features that are absent from local pseudopotential calculations.

2. Since the s part of our nonlocal potential is substantially deeper than the local pseudopotentials used previously,\(^15\) (Fig. 1), the lowest (As-s-like) bulk valence band in GaAs as well as the lowest As-derived surface states As(1) and As(2) (Fig. 2) are pulled down in energy by about 1.5 eV, in agreement with recent experimental results.\(^25\)\(^-\)\(^30\) Tight-binding calculations\(^13\)\(^-\)\(^23\) place the As(2) level in resonance with the bulk bands, whereas local-pseudopotential calculations\(^15\) place the As(2) state correctly in the heteropolar gap, but too high by about 1.3 eV.

3. Whereas a local-pseudopotential calculation predicts only one clearly defined As-p-like state in the (−2.4) eV region, the nonlocal calculation predicts two states, As(3a) and As(3b), at (3.0–3.5) eV with nearly identical dispersion, in agreement with experiment.\(^25\)\(^-\)\(^30\) These states are relaxation induced and are localized on the second surface layer. Hence, the nonlocal pseudopotentials respond more sensitively to changes in the atomic positions and split two rather than one surface As-p state from the bands.

4. Both As and Ga atoms have relatively low-lying 4d states that are unoccupied in the ground atomic state but may mix into the bands in the bulk solid and surface. Whereas the d component
of the nonlocal pseudopotential is appreciably deeper than the s component for both Ga and As (Fig. 1) (reflecting a weaker pseudopotential cancellation due to the existence of only one core state of d symmetry compared to three of s symmetry), the local pseudopotential approach constrains both s and d states to sample a common potential. This raises artificially the energy of the d-containing states to the higher conduction bands, outside the energy region of conventional interest. Hence, in this approach the valence, surface, and lowest conduction states are characterized as s- and p-like states. In contrast, an s, p, d nonlocal pseudopotential shows As d character in the lowest unoccupied surface state Ga(3) (cf. Fig. 2) as well as Ga d character in the highest occupied As(5) surface state. This may lead to a possible reinterpretation of some of the selection rules used to analyze the partial-yield photoemission data.31

(5) Angular-resolved photoemission studies in which the emission was collected in the mirror plane have indicated three or four surface states around X' at -0.9, -1.4, and -1.6 eV below the valence-band edge.30 The two outer states at -0.9 and -1.6 eV were observed to be strong for a polarization parallel to the mirror plane (i.e., even with respect to reflection), whereas the central state at -1.4 eV is strong only for a polarization direction perpendicular to the mirror plane (i.e., odd). In contrast with the local pseudopotential study,15 which finds an order of surface states of odd, even, even [As(6), As(5), As(4), respectively], I find the correct order of even, odd, even [As(5), As(3), As(4)] not as X’ but rather at a point displaced by 20% towards X. I do not find the As(6) state that was obtained in the local pseudopotential study15 (localized parallel to the surface plane in the X'-X' region of the bands). A correct order of states in this region was found also in a tight-binding study13 but for a somewhat different surface geometry (bond-rotation relaxation rather than the presently used 27°-rotation relaxation model). In agreement with recent experimental studies,23 I find four surface states between X' and Γ above -3 eV, whereas the local-pseudopotential study indicates only three states in this region. The predicted dispersion of the four states along X' to Γ agrees very well with experiment25 and suggests an assignment that differs from that suggested both by tight-binding23 and the local-pseudopotential15 studies.

II. METHOD

A self-consistent symmetrized plane-wave approach in a repeated slab geometry with nine layers of GaAs (i.e., 18 atoms) and six layers of vacuum in each of the translationally equivalent unit cells is used here. The effective potential in coordinate space is partitioned into an external pseudopotential W_{ext}(\mathbf{r}, \mathbf{r}') and a two-electron screening potential W_{se}(\mathbf{n}(\mathbf{r})) that is a functional of the self-consistent pseudodensity n(\mathbf{r}):

\[ V_{\text{eff}}(\mathbf{r}, \mathbf{r}') = W_{\text{ext}}(\mathbf{r}, \mathbf{r}') + W_{\text{se}}[n(\mathbf{r})]. \]

(1)

The pseudopotential acts as an external field, replacing the dynamic effects of the core electrons, and is a constant for a given geometry of the system. The nonlinear response of the electronic system to this external potential is represented by the screening, which is calculated self-consistently. Note that whereas certain choices of screening formalisms may lead to a local screening potential W_{se}(\mathbf{r}) (e.g., the local-density approach5) and others (e.g., the Hartree–Fock approach) lead to a nonlocal screening potential W_{se}(\mathbf{r}, \mathbf{r}') both approaches produce nonlocal pseudopotentials W_{\text{ext}}(\mathbf{r}, \mathbf{r}') : eliminating a certain subspace of (core) orbitals from either a local or a nonlocal all-electron representation requires the addition of a state-dependent (pseudo) potential.14

The pseudopotential W_{\text{ext}}(\mathbf{r}, \mathbf{r}') is approximated as a superposition of angular-momentum-dependent atomic pseudopotentials v^{(\alpha l)}_{\text{ext}}(\mathbf{r}) for each atomic type \alpha:

\[ W_{\text{ext}}(\mathbf{r}, \mathbf{r}') = \sum_{\mathbf{R}_n} \sum_{\mathbf{T}_\alpha} \left[ v^{(\alpha l)}_{\text{ext}}(\mathbf{r} - \mathbf{R}_n - \mathbf{r}_\alpha) + v^{(\alpha l)}_{\text{ext}}(\mathbf{r}' - \mathbf{R}_n - \mathbf{r}'_\alpha) - v^{(\alpha l)}_{\text{ext}}(\mathbf{r} - \mathbf{R}_n - \mathbf{r}_\alpha) - v^{(\alpha l)}_{\text{ext}}(\mathbf{r}' - \mathbf{R}_n - \mathbf{r}'_\alpha) \right] \hat{P}_l. \]

(2a)

I use v(\mathbf{r}) to indicate atomic potentials; V(\mathbf{r}) denotes crystal potentials. Here \mathbf{R}_n and \mathbf{r}_\alpha denote the unit-cell and atomic-site-position vectors, respectively, and \hat{P}_l is an angular momentum projection operator with origin at \mathbf{r} - \mathbf{R}_n - \mathbf{r}_\alpha. The total pseudopotential has been divided into a local part V_{\text{loc}}(\mathbf{r}) common to all angular momentum components of the wave function and into a nonlocal part that acts differently on the different angular components of the wave function. Note that the definition of the local atomic potential v^{(\alpha l)}_{\text{loc}}(\mathbf{r}) is arbitrary as long as the nonlocal part V_{\text{non}}(\mathbf{r}) includes all the differences v^{(\alpha l)}_{\text{ext}}(\mathbf{r} - \mathbf{R}_n - \mathbf{r}_\alpha) as indicated in Eq. (2). Since at a large distance from the origin all atomic pseudopotentials approach -Z/\mathbf{r}, the nonlocal potential is nonzero predominantly inside the core. One is hence free to select v^{(\alpha l)}_{\text{loc}}(\mathbf{r}) = A_{\alpha}^{(0)}, A_{\alpha}^{(1)}, A_{\alpha}^{(2)} + A_{\alpha}^{(0)}, A_{\alpha}^{(1)}, A_{\alpha}^{(2)} and optimize A_{\alpha}, A_{\alpha}, and A_{\alpha} to obtain a good convergence in the basis-set expansion.
Similarly, one may select \( V^c (\mathbf{r}) \) to be equal to the local semiempirical pseudopotential \( V^{\text{semp}} (\mathbf{r}) \) used in previous studies\(^5\) and define the nonlocal part as the deviation of the first-principles pseudopotential from \( V^{\text{semp}} (\mathbf{r}) \):

\[
W_{\text{ps}}(\mathbf{r}, \mathbf{r'}) = \sum_{\mathbf{k} \mathbf{a}} v_{\text{ps}}^{(a)}(\mathbf{r} - \mathbf{R}_a - \mathbf{r}_a) + \sum_{\mathbf{k} \mathbf{a}} \sum_{\mathbf{r}} v_{\text{ps}}^{(a1)}(\mathbf{r} - \mathbf{R}_a - \mathbf{r}_a) - v_{\text{ps}}^{(a)}(\mathbf{r'} - \mathbf{R}_a - \mathbf{r}_a) \mathbf{P}_a .
\]  

(2b)

In the momentum representation used here, the matrix element of the pseudopotential between two plane-wave basis functions \( \exp[i(\mathbf{k} + \mathbf{G}) \cdot \mathbf{r}] \) and \( \exp[i(\mathbf{k} + \mathbf{G'}) \cdot \mathbf{r}'] \) is

\[
\langle \mathbf{k} + \mathbf{G} | W_{\text{ps}}(\mathbf{r}, \mathbf{r'}) | \mathbf{k} + \mathbf{G'} \rangle = \sum_a S_a(\mathbf{G}) F_{a \sigma \bar{\sigma}}(\mathbf{G}) + \sum_a \sum_{\mathbf{G}} S_a(\mathbf{G} - \mathbf{G'}) F_{a \sigma \bar{\sigma}}(\mathbf{G} + \mathbf{G'}) ,
\]  

(3)

where the atomic nonlocality matrix in momentum space is given by

\[
F_{a \sigma \bar{\sigma}}(\mathbf{G} + \mathbf{G'}) = \langle \mathbf{k} + \mathbf{G} | v_{\text{ps}}^{(a \bar{\sigma})}(\mathbf{r}) - v_{\text{ps}}^{(a \sigma)}(\mathbf{r}) | \mathbf{k} + \mathbf{G'} \rangle = (4\pi/\Omega_g)(\mathbf{1} + 1)(j_1(|\mathbf{k} + \mathbf{G} | |\mathbf{r} |)[v_{\text{ps}}^{(a \bar{\sigma})}(\mathbf{r}) - v_{\text{ps}}^{(a \sigma)}(\mathbf{r})]) P(\cos \gamma) |j_1(|\mathbf{k} + \mathbf{G'} | |\mathbf{r} |) .
\]  

(4)

where \( \Omega \) is the superlattice unit-cell volume.

In a local pseudopotential approximation\(^5\) one sets

\[
F_{a \sigma \bar{\sigma}}(\mathbf{k} + \mathbf{G} + \mathbf{G'}) = 0
\]  

(11)

for all angular momentum components. Hence, the \( s, p, d, \) etc., components of the wave function are all constrained to sample a common local potential. This is a good approximation only to the extent that (i) \( v_{\text{ps}}^{(a \bar{\sigma})}(\mathbf{r}) \approx v_{\text{ps}}^{(a \sigma)}(\mathbf{r}) \approx v_{\text{ps}}^{(a \sigma)}(\mathbf{r}) \), or if (ii) \( v_{\text{ps}}^{(a \bar{\sigma})}(\mathbf{r}) \neq v_{\text{ps}}^{(a \sigma)}(\mathbf{r}) \neq v_{\text{ps}}^{(a \sigma)}(\mathbf{r}) \), but that the crystal wave functions have only one predominant angular momentum component over the entire energy-band range of interest. In practice, neither of these conditions are met for covalently bonded tetrahedral solids that exhibit extensive orbital hybridization over the full valence and lowest conduction bands [cf. condition (ii)] and are made of atoms with a different number of core orbitals for different \( l \) values [cf. condition (i)].

In the present calculation, I expand the pseudowave-functions \( x_j \) of band \( j \) and momentum \( \mathbf{k} \) in a plane-wave basis:

\[
x_j(\mathbf{k}, \mathbf{r}) = \sum_{\mathbf{d}} B_j(\mathbf{k} + \mathbf{G}) e^{i(\mathbf{k} \cdot \mathbf{r})} ,
\]  

(12)

followed by a symmetrization of \( \{ e^{i(\mathbf{k} \cdot \mathbf{r})} \} \) with respect to a reflection operation at the center of the unit cell. This symmetry operation belongs to the star of all \( \mathbf{k} \) points in the Brillouin zone of the \((110)\) system and reduces substantially the matrix size. With the basis of the form (12), the single-particle equation becomes

\[
\sum_{\mathbf{d}} [ (\mathbf{k} + \mathbf{G})^2 \delta_{\mathbf{d}, \mathbf{d'}} + W(\mathbf{k} + \mathbf{G}, \mathbf{k} + \mathbf{G}) - \epsilon_j(\mathbf{k}) ] B_j(\mathbf{k} + \mathbf{G}) = 0 ,
\]  

(13)

where the potential matrix is given in the momentum representation as
\[ V(\mathbf{k} + \mathbf{G}, \mathbf{k} + \mathbf{G}') = V_{\text{se}}(\mathbf{G} - \mathbf{G}') + V_{\text{c}}(\mathbf{G} - \mathbf{G}') + V_{\text{c}}(\mathbf{G} - \mathbf{G}') \]
\[ + \sum_a S_a(\mathbf{G} - \mathbf{G}') \left( F_{\alpha 0}(\mathbf{G} - \mathbf{G}') + \sum F_{\alpha l}(\mathbf{k} + \mathbf{G}, \mathbf{k} + \mathbf{G}') \right), \]

and the various terms are given by Eqs. (10a), (10b), (10c), (6), and (4), respectively.

The input to the calculation is \( \{ v_{\text{se}}^{(a)}(r) \} \) for \( a = \text{Ga, As} \) and \( l = 0, 1, 2 \), as well as the geometry \( \{ \mathbf{R}_a \}, \{ \tau_a \} \).

The matrix elements \( F_{\alpha 0}(\mathbf{G}) \) and \( F_{\alpha l}(\mathbf{k} + \mathbf{G}, \mathbf{k} + \mathbf{G}') \) are then calculated once and for all on a given grid in momentum space by performing the one-dimensional numerical integrations indicated in Eqs. (6) and (4), respectively. To solve (13), an initial guess for \( V_{\text{se}}(\mathbf{G}) + V_{\text{c}}(\mathbf{G}) + V_{\text{c}}(\mathbf{G})' \) is needed. This guess does not affect the final self-consistent result but rather the computing speed with which it is obtained. One can use an initial guess taken from previous empirical pseudopotential studies or by first solving an atomic pseudopotential equation with \( v_{\text{se}}^{(a)}(r) \) for \( \text{Ga and As} \):

\[ \{- \frac{1}{2} \nabla^2 + v_{\text{se}}[n_{\alpha}(r)] + v_{\text{c}}[n_{\alpha}(r)] + v_{\text{c}}'[n_{\alpha}(r)] + v_{\text{se}}^{(a)}(r) \} \psi_{\alpha\text{el}}(r) = \epsilon_{\alpha\text{el}} \psi_{\alpha\text{el}}(r) \]

using standard atomic structure programs, and then use the linearly superposed atomic screening as a first guess to the crystal screening:

\[ W_{\text{sec}}^{(a)}(\mathbf{r}) = \sum_{\mathbf{k}} \sum_{\tau_a} \{ \psi_{\text{sec}, \tau_a}(n_{\alpha}(\mathbf{r} - \mathbf{k} - \tau_a)) \}, \]

where \( \psi_{\text{sec}, \tau_a}(\mathbf{r}) \) denotes collectively \( \psi_{\text{se}} + \psi_{\text{c}} + \psi_{\text{c}}' \). Given this initial guess, Eq. (13) is solved for four special \( \mathbf{k} \) points in the surface Brillouin zone. To obtain a convergence of 0.25 eV in the eigenvalues \( \epsilon_{\mathbf{k}}(\mathbf{G}) \) with respect to the sum over \( \mathbf{G} \) in Eq. (13), I use 600 plane waves at \( \Gamma \) (which decompose into two 300 \times 300 matrices due to the symmetrization) plus an additional 1167 plane waves in a second-order Lowdin perturbation technique. These correspond to energy cutoff values of 4.1 and 9.5 Ry, respectively. Those large cutoff values are required because the first-principles pseudopotentials (Fig. 1) are rather steep in coordinate space.

From the wave functions \( x_{\mathbf{k}}(\mathbf{r}, \mathbf{r}) \) evaluated at points \( \mathbf{k}_p \) in the irreducible zone [Eq. (12)] one calculates the symmetrized wave function \( \tilde{x}_{\mathbf{k}}(\mathbf{r}, \mathbf{r}) \) over the full zone. The charge density is then given as

\[ n(\mathbf{r}) = \sum_{\mathbf{k}} \sum_{\tau_a} \sum_{\mathbf{k}_p} w(\mathbf{k}_p) \tilde{x}^*_{\mathbf{k}_p}(\mathbf{r}, \mathbf{r}) \tilde{x}_{\mathbf{k}_p}(\mathbf{r}, \mathbf{r}), \]

using the statistical weights \( w(\mathbf{k}_p) \) for the four special \( \mathbf{k} \) points \( \mathbf{k}_p \). The density Fourier components are

\[ n(\mathbf{G}) = \sum_{\mathbf{k}} \sum_{\tau_a} \sum_{\mathbf{k}_p} w(\mathbf{k}_p) B^*_{\mathbf{k}_p}(\mathbf{G}) B_{\mathbf{k}_p}(\mathbf{G} + \mathbf{G}'). \]

From (18) I obtain the interelectronic Coulomb potential \( V_{\text{se}}(\mathbf{G}) \) using Eq. (10a). The individually divergent \( V_{\text{se}}(\mathbf{G} = 0) \) and \( F_{\alpha 0}(\mathbf{G} = 0) \) terms are arbitrarily set to zero, redefining thereby the position of the vacuum level. From the coordinate space charge density in Eq. (17), tabulated over about 65,000 grid points \( \{ \mathbf{r}_l \} \) in the unit cell, I calculate \( n(\mathbf{r}) )^{1/3} \), and via a fast Fourier transform, I obtain from this the \( V_{\text{c}}(\mathbf{G}) \) term in (10b).

The correlation potential (10c) is similarly calculated by fast Fourier transforms of \( V_{\text{c}}(\mathbf{k}) \). Using this updated screening \( W_{\text{sec}} = V_{\text{se}}(\mathbf{G}) + V_{\text{c}}(\mathbf{G}) + V_{\text{c}}(\mathbf{G}) \), and the fixed pseudopotentials \( F_{\alpha 0}(\mathbf{G}) \) and \( F_{\alpha l}(\mathbf{k} + \mathbf{G}, \mathbf{k} + \mathbf{G}') \), Eq. (13) is solved again. The process is repeated iteratively until the screening \( W_{\text{sec}}(\mathbf{G}) \) agrees to within \( 10^{-4} \) Ry in successive iterations. Starting from the initial guess (16) and using extensively iteration-damping techniques, this requires nine iterations.

III. PSEUDOPOTENTIALS

Our basic method for obtaining nonempirical atomic pseudopotentials \( \{ v_{\text{se}}^{(a)}(r) \} \) in the density-functional formalism has been previously described. The nonlocal density-functional pseudopotentials for Ga and As are displayed in Fig. 1, where they are compared with the semiempirical local pseudopotentials used previously for surface calculations. These latter potentials are obtained by fitting the spectra of the bare ions \( \text{Ga}^+ \) and \( \text{As}^+ \) to experiment without constraining the shape of the wave functions. For comparison, Fig. 1 also shows the local pseudopotential developed by Fresnley and Kroemer by fitting orbital energies and using realistic Hartree-Fock-type atomic charge densities. It is seen that the semiempirical local pseudopotential is qualitatively different from the two other potentials which incorporate realistic charge densities: Its minimum is displaced to larger \( \mathbf{r} \) values and is substantially shallower. Having also an attractive character in the core region, the semiempirical pseudopotential tends to accumulate charge in the core (where a pseudopotential description of the wave function is least valid). Due to the wave-function normalization requirement, such a core
attraction tends to deplete charge density from the bond region. Although the density-functional $l = 0$ pseudopotentials are very similar to the local FK pseudopotentials (used by these authors also to obtain a good fit of the GaAs band structure to experiment\textsuperscript{35}), the former have, in addition, $l = 1,2$ components. The $l = 1,2$ ($p$-like and $d$-like, respectively) density-functional pseudopotentials are appreciably stronger than the $l = 0$ ($s$-like) pseudopotential. Application of the local semi-empirical pseudopotentials to the calculation of the ground state of Ga and As atoms indeed indicates that the $p$ orbitals lie at about 1.3 eV above the values obtained both with the “full” potential and the density-functional pseudopotential. Similarly, the atomic orbitals outside the core differ considerably (by about 15\%) from the “true” all-electron valence orbitals. Note that a significant difference persists between the $v_0^{s(0)}$ potential and $v_0^{s(1)}$, $v_0^{s(2)}$, even after addition of the $1/r^2$ and $3/r^2$ centrifugal terms to $v_0^{s(1)}$ and $v_0^{s(2)}$, respectively.

While it is conceivable that a suitably local pseudopotential for Ga and As may be constructed to approximate the results obtained with the full nonlocal pseudopotentials, no such simple potential exists as yet. Not only does the semi-empirical local pseudopotential\textsuperscript{15} fail to reproduce well the results obtained with the nonlocal pseudopotentials, but also the present $v_0^{s(0)}$ alone fails in providing a good approximation to the nonlocal results.

The inclusion of pseudopotential nonlocality has significant effects on the character of the wave functions. Although atomic As and Ga do not have occupied 4$d$ states in the ground state, their s, p valence state may hybridize with d states in the solid. Indeed, the calculated 4$p$-to-4$d$ one-electron gap in the atoms (8.5 and 5.5 eV for As and Ga, respectively) is of similar magnitude to the 4$s$-to-4$p$ valence gap (9.5 and 7 eV for As and Ga, respectively). Furthermore, the 4$d$ atomic orbital of Ga overlaps significantly with the 4$s$, 4$p$ orbitals of As when these atom are displaced by a typical Ga-As bond length. Such an $s$-$d$ and $p$-$d$ hybridization is made unlikely if one uses local pseudopotentials, since the $d$-like components of the crystal wave functions are forced to sample a shallow local pseudopotential, more akin to s states, rather than the deeper nonlocal d pseudopotential (cf. Fig. 1). This pushes the $d$-containing crystal states upwards in energy into the higher conduction bands, outside the energy region of physical interest. The d nonlocality has also an indirect effect on the s states: these can penetrate the core region, which lowers their energy considerably.\textsuperscript{25} A local pseudopotential approximation may hence be too crude to reveal reliably the propor-

FIG. 1. Nonlocal density-functional atomic pseudopotentials compared with the local semi-empirical pseudopotential (Ref. 15) and the pseudopotentials of Frensley and Kroemer (Ref. 38): (a) Ga, (b) As.
tions of s, p, and d character in each state, while energies that are stationary with respect to small wave-function variations may be better represented.

A note is in order here on a related feature of the tight-binding (TB) method. As core states are omitted from explicit consideration by this approach, it constitutes in effect an implicit pseudopotential method. However, contrary to the explicit pseudopotential approaches, the basis functions remain unspecified, and (in current versions of TB) they are constrained to have zero intersite overlap and possess matrix elements with s and p symmetries alone. Hence this excludes the possibility of having direct d character in the bands at any energy, while local pseudopotentials merely tend to raise d-containing states to higher energies.

The density-functional nonlocal pseudopotentials have been smoothed in the inner core for \( r < 0.5 \) a.u. to avoid numerical instabilities resulting from the potential fluctuations of this region. The criteria used for smoothing is that the smoothed potential equals approximately \( \int_0^{\text{max}} \delta_\alpha(q) e^{iqr} dq \) with \( \text{max} \) being the highest Fourier component included in the Hamiltonian matrix. This has a small effect on the atomic structure (the energy eigenvalues of the smoothed potential differ by less than 0.01 eV from those obtained by the unsmoothed potential) and enables better numerical convergence. Applying these potentials for bulk GaAs using \( I = 0, 1, 2 \) for both atoms, a reasonable band structure is obtained. The energy eigenvalues for high-symmetry points (the value, in eV, calculated with the first-principles nonlocal pseudopotential is given first, followed by the value obtained by Chelikowsky and Cohen) by fitting the experimental data (this, however, is not the potential used by these authors for surface calculations) and the experimental value, when available] are: \( \Gamma_{1s} (\sim -13.7), -12.6, -13.8 \pm 0.4) \), \( \Gamma_{1p} (1.0, 1.51, 1.52) \), \( \Gamma_{3p} (3.8, 4.6), X_{1s} (-11.0, -9.8, -10.7 \pm 0.3) \), \( X_{3s} (-7.2, -6.9, -7.1 \pm 0.2), X_{3p} (-2.5, -2.9), X_{1s} (1.7, 2.03), X_{3s} (1.3, -1.3, -1.4 \pm 0.3) \). No comparison can be made with the band structure obtained with the local pseudopotential of Ref. 15, since the results were not published. The main defect of the density-functional band structure lies in the small predicted band gap (1.0 eV compared with the observed value of 1.52 eV). This is by now a well-recognized general shortcoming of the density-functional approach for screening, both in its pseudopotential and all-electron forms (e.g., for Si the band gap calculated either with the full potential or with the first-principles nonlocal pseudopotential is 0.5–0.6 eV, compared with the observed value of 1.17 eV, whereas for CdS it is 2.0 eV, compared with the experimental value of 2.55 eV).

The main sources of error in our calculation are:

1. In principle, one should also include pseudopotential components with \( I > 2 \). However, using \( I = 0, 1, 2 \) alone is probably a good approximation for semiconductors, since wave-function components with \( f \) or \( g \) character are not expected to be pronounced in the valence and the low conduction bands.

2. Using a repeated slab geometry introduces spurious interactions between surface states on opposite sides of the slab. If one does not force the degeneracy of the energy eigenvalues resulting from this symmetry, the splitting between them forms a measure of the spurious intersurface interactions. The off-degeneracy is found to be 0.2–0.3 eV. Since the density-functional pseudopotentials are more localized than the local semiparametric pseudopotentials (Fig. 1), one expects such spurious interactions to be smaller in the present study.

3. The most serious source of numerical error in the present calculation involves the truncation of the Fourier expansions at finite momentum \( q \). One can estimate this error by changing the momentum cutoff values and repeating the band calculation. I use two values of the cutoff parameter \( q = 0.9q_0 \) and \( 1.1q_0 \), where \( q_0 \) is the largest wave vector entering the Hamiltonian matrix. For each such \( q \) value one defines a corresponding smoothed atomic pseudopotential \( \int_0^\infty \delta_\alpha(q) e^{iqr} dq \) where \( \delta_\alpha(q) \) is the Fourier transform of the untruncated nonlocal atomic pseudopotential. Using such smoothed pseudopotentials one is assured that the solution of the eigenvalue problem (Eq. (13)) with \( G = q \) reflects the true results of the assumed potential. I find that the average deviation of the band energies from those obtained using \( q_0 \) as a cutoff, over an energy range of 16 eV of valence and conduction states, is 0.32 eV for \( q = 0.9q_0 \) and 0.18 eV for \( q = 1.1q_0 \).

IV. GEOMETRY

The present study uses the surface geometry deduced by Tong et al.12 from LEED studies. This relaxation model involves an angle of 27° in the inward rotation of Ga relative to the "ideal" surface plane. Only surface atoms are allowed to relax; in units of the bulk interplanar distance \( d_0 \), the As atoms move outwards by about 0.1 \( d_0 \), and the Ga atoms move inwards by about 0.25 \( d_0 \). A complete discussion of the various surface geometry models inferred from LEED studies is given in Refs. 44 and 45.
V. RESULTS

The (110) projected band structure of GaAs together with the calculated surface bands are displayed in Fig. 2(a). A schematic drawing of the experimentally observed surface states$^{22,23,29,30}$ plotted with the calculated projected band structure is given in Fig. 2(b). The surface states have been denoted by the chemical symbol of the atom (Ga/As) that forms the predominant orbital character in the corresponding wave function, followed by a number in parentheses, in increasing order from the bottom of the bands that labels the different states. Fifteen calculated \( \mathbf{k} \) points in the surface Brillouin zone have been used, and the wave functions of the lowest 60 bands at each point have been analyzed by calculating the wave function's planar average (which indicates the region of space perpendicular to the surface where the wave function is mostly localized) as well as an angular momentum projection of the wave function \( \langle \mathbf{r}^2 \rangle \mathbf{Y}_{lm} \) around spheres (with \( r = R_s \) where \( R_s \) is the tetrahedral radius$^{40}$) centered on surface atoms. This can be used to establish the predominant orbital character in each state, and when normalized, this yields the percentage of a given angular momentum species in each state. As both \( s \) as well as \( p, d \)-type pseudopotentials wave functions have zero amplitude at the origin, such a decomposition is an important tool in analyzing the orbital character. In addition, a large number of the calculated wave functions have been subjected to a symmetry reflection operation in the mirror plane to examine their parity.

The two lowest surface states labeled As(1) and As(2) \( (S_2 \text{ and } B_3, \text{ respectively, in the notation of Ref. 13 and Ref. 16}) \) are As \( s \)-like. The As(1) state was observed experimentally at \(-12\, \text{eV (Ref. 30)}\) and is predicted to have its region of maximum band flatness (i.e., peak density) at \(-12.1\, \text{eV} \) by the present theory and at \(-10.2\, \text{eV} \) by the local semiempirical pseudopotential theory.$^{15}$ The As(2) state has been observed at \(-11\, \text{eV (Ref. 30)} \) (it is, however, not clear whether the emission is from a surface or a bulk state) and is predicted by the present theory to lie at \(-10.8\, \text{eV (X)} \) and \(-10.7\, \text{eV (M)} \), whereas the local pseudopotential theory places it at \(-9.3\, \text{eV} \).$^{15}$

Tight-binding calculations seem to place the As(2) state in the bulk bands. Whereas the As(1) is localized predominantly on the plane next to the surface, the As(2) state is strongly localized on the surface plane. The discrepancies between the local$^{15}$ and nonlocal pseudopotential results for these states occur because of the misplacement of the low As-derived bulk band

FIG. 2. Projected surface-band structure of GaAs (110): (a) nonlocal-pseudopotential calculation. Full lines and dashed lines indicate strong and weak surface states, respectively; (b) schematic representation of the experimentally observed surface states (Refs. 20–23, 29, 30) plotted with the calculated projected bulk band structure.
by the former method. More detailed experimental data are needed to assess the calculated dispersion and energy separation between these states.

The Ga(1) and Ga(2) pair of bands appearing at $-6.5\text{ eV}$ at $\overline{M}$ (in the notation of Ref. 13) are Ga s-like with a 10% As $p$-like contribution. Similarly to the As(1)–As(2) pair, Ga(1) is localized on the second layer, whereas Ga(2) is localized on the first layer. Only one state at $-6.5\text{ eV}$ (Ref. 23) is observed experimentally in this energy region. Both pairs As(1)–As(2) and Ga(1)–Ga(2) are split by an interaction induced by the surface relaxation (1.5 and 0.3 eV, respectively).

The As(3a) and As(3b) pair of surface states are localized on the two upper surface layers and are As $p$-like with 25% admixture of Ga $p$. Two surface states are observed experimentally in this energy region at $-3.6$ to $-3.8\text{ eV}$ ($X$ to $\overline{M}$) and $-3$ to $-3.2\text{ eV}$ ($X' \rightarrow M$). The calculated positions are As(3a) at $-3$ to $-3.2\text{ eV}$ and As(3b) at $-2.9\text{ eV}$. The local-pseudopotential theory predicts one state in this region at $-3.2\text{ eV}$ (plus an additional very weak state near the zone corner). Here it is found that the splitting of these states is small (0.1–0.2 eV), reflecting a different content of the minority character (Ga $p$). With respect to the symmetry operation of mirror–plane reflection, I find As(3b) to be odd and As(3a) to be even along the $\overline{M} \rightarrow X'$ line.

The As(5) and As(4) pair of states ($B_1$ and $S_2$ in the notation of Ref. 13 or the As “dangling” and As “back” bonds, respectively) are localized on the first surface layer. Whereas they are classified as As $p$-type both in the tight-binding and local-pseudopotential calculations, I find these states to include also a non-negligible As $d$ character (e.g., 30% in As(5) and 17% in As(4) at the $\overline{M}$ point with 75% and 80% $p$ character, respectively). I find the As(4) and As(5) states to be even under a mirror–plane reflection at $\overline{X}'$ and $\overline{M}$; however, the As(4) state reverses its parity to odd just outside $X'$ towards $\overline{M}$, and As(5) continues to be even under reflection. The angular-resolved photoemission study of Williams, Smith, and Lapere38 shows that at this energy level at least three states at $-0.8$, $-1.4$, and $-1.6\text{ eV}$. At $\overline{X}'$ I find the order As(5) at $-0.8\text{ eV}$, As(4) at $-1.2\text{ eV}$, As(3b) at $-1.65\text{ eV}$, and As(3a) at $-1.7\text{ eV}$. With respect to mirror reflection, these states are even, odd, even, respectively. If one moves, however, away from $X'$ towards $\overline{M}$, the order is changed to As(5), As(3b), As(4), and As(3a) with polarities even, odd, even, respectively. Williams et al.38 have studied the angular-resolved energy distribution curves obtained with a polarization of the incoming radiation both parallel and perpendicular to the mirror plane and collecting the emission in the mirror plane. They concluded that the states at $-0.9$ and $-1.6\text{ eV}$ are strong for a polarization parallel to the mirror plane (even), whereas the transition at $-1.4\text{ eV}$ is odd, i.e., the order of states of increasing binding energies is even, odd, even, respectively. This agrees with the order found here at a point that is 20% away from $X'$ towards $\overline{M}$ but disagrees with the order predicted at $X'$. Huijser et al.23 have shown four states near $X'$, the two lowest being nearly degenerate at $X'$ but splitting appreciably in going from $X'$ to $\overline{M}$. The observed order of the dispersion of these states in going from $X'$ to $\overline{M}$ (and from the least bound to the most tightly bound) is\textsuperscript{25}: up, up, dispersionless, and down. These observations strongly support the ordering As(5), As(3b), As(4), and As(3a), which agrees with experiment\textsuperscript{23} both in the dispersion directions as well as with the assignment of polarities\textsuperscript{26} even, odd, odd, even. This interpretation is different from the one suggested by the tight-binding work\textsuperscript{13} [As(5), As(4), As(3), or $B_1$, $S_1$, and $S_2$ respectively, where As(4) is found to be odd] which agrees with experiment in the order of polarities but shows only three rather than four surface states at $\overline{M} \rightarrow X'$ above $-3\text{ eV}$. Similarly, the present interpretation differs from the one suggested by the local-pseudopotential work, which shows only three states [As(6), As(5), As(4) or odd, even, even, where As(4) is even in agreement with the present results but in disagreement with Ref. 13]. I do not observe any As(6) state (in agreement with Ref. 13), but instead the As(2) state extends to $X'$ (in agreement with Ref. 13). In order to verify whether the order parity and dispersion of the narrowly spaced bands near $X'$ depend strongly on the number of plane waves used in the calculation, two calculations were repeated using a basis set which includes $+10\%$ and $-10\%$ more (less) basis functions. I find that the convergence error near $X'$ is mostly $0.12\text{ eV}$ rigid (i.e., state independent) with a small $0.04\text{ eV}$ nonrigid component which does not alter the order, parity, and dispersion direction of the bands.

In the previous interpretations of the polarities of the surface states,\textsuperscript{12–23} it was simplistically assumed that the initial states are of $p$ symmetry and the final states are of $s$ symmetry. In the present study, non-negligible $d$ character is found in both states. In particular, the lowest empty surface state Ga(3) includes $p$ symmetry on the Ga site (91% at $X$) but 35% $d_x$ character on the As site (as well as 30% As $s$ on As); note that bonding to an electronegative chemisorbed species will tend to stabilize such an orbital. Tight-binding calculations\textsuperscript{23} exclude a direct $d$ character from
the problem by avoiding the use of \( d \) basis functions and neglecting intersite overlap. As indicated before, whereas local-pseudopotential studies allow, in principle, \( d \) character in the wave functions (as the plane-wave basis used can produce all angular symmetries), the neglect of nonlocality effects tends to destabilize \( d \) states and push them to high energies.

The empty surface states \( \text{Ga}(3) \) and \( \text{Ga}(4) \) obtained in this study are localized on the surface layer and agree closely in location with the corresponding states found in the local-pseudopotential study.\(^{15} \) As indicated by others,\(^{13,15} \) the \( \text{Ga}(4) \) state lies in the gap for the unrelaxed geometry, whereas \( \text{Ga}(3) \) lies above it. Relaxation effects reverse their order and place both of them in the conduction band. Partial-yield photoemission studies\(^ {25,31} \) indicate that transitions from \( \text{As} 3p \) and \( \text{Ga} 3d \) to empty surface states have comparable intensities, whereas the \( \text{As} 3d \)-to-empty-state transitions are extremely weak. It was assumed\(^ {31} \) that if cross transitions coupling the two sublattices are weak, the \( \text{As} \) character in the empty states is \( s \)-like and the \( \text{Ga} \) character is \( p \)-like, in agreement with the present study. However, the absence of transitions from \( \text{As} 3d \) is also naturally explained by the fact that the present study finds \( As \) \( d \) character in the empty states.

In regard to the assignment of certain atomic orbital character to surface or bulk states, one notes that such procedures involve a certain amount of arbitrariness: a three-dimensional charge density can not be partitioned in any unique way into its atomic parentage whether one uses localized\(^ {13,14} \) or delocalized\(^ {15} \) basis functions. The nonlocal-pseudopotential approach offers, however, a physical way of assessing the orbital character of surface states simply by scaling down by a small amount the nonlocal-pseudopotential component \([\epsilon_{\text{np}}(\text{As}) - \epsilon_{\text{np}}(\text{Ga})] \) \([\text{Eq. (2b)}\) of a given \( l \) for a chosen atomic species \( \alpha \). To estimate the nonlocality effects, I have repeated a non-self-consistent calculation \([i.e., \text{using the screening, which is self-consistent, with the full nonlocal pseudopotential in Eq. (2b)}\) with a perturbational change in the nonlocality matrix elements \( F_{\alpha l}(\mathbf{k} + \mathbf{G}, \mathbf{k} + \mathbf{G}') \rightarrow \lambda F_{\alpha l}(\mathbf{k} + \mathbf{G}, \mathbf{k} + \mathbf{G}') \) with \( \lambda = 1 \pm 0.05 \). Whereas the choice \( \lambda = 0 \) reduces the problem to a local-pseudopotential calculation, the choice \( \lambda = 1 \pm \eta \) with a small \( \eta \) allows one to examine how the electronic charge and orbital energies of various surface states respond to small perturbations in the external nonlocal potential, revealing thereby their orbital character. Note that while the total elimination of the nonlocal components \( (\lambda \rightarrow 0) \) will yield results that are determined by the choice of the local pseudopotential, the use of a limiting procedure \( (\lambda = 1 \pm \eta \) with \( \eta \ll 1) \) allows one to examine the derivative of the energy of a given surface state with respect to the potential nonlocality. To first order, this derivative is independent of the local potential.

First, the \( \text{Ga} \) \( d \) nonlocality was scaled down. The major effect was to raise the energy of the empty states \( \text{Ga}(3) \) and \( \text{Ga}(4) \) with \( d \epsilon / d \lambda \approx 0.15 \) eV and to decrease the \( d \) character of the wave function by increasing substantially the \( s \) character \([e.g., \text{from 9}\% \text{to 16}\% \text{in} \text{Ga}(3) \text{at the limit} \lambda \rightarrow 0] \). Although self-consistency will tend to somewhat offset this linear response, it seems clear that the \( \text{Ga} \) \( d \) character has an important role in stabilizing the empty \( \text{Ga} \) states.

Among the other choices of nonlocal pseudopotentials that have been scaled, an interesting result is obtained when the \( s, p \) nonlocality is smoothly scaled towards zero \((i.e., s, p \) electrons tend to feed a common local potential). It is found here that the splitting between the \( \text{As}(3a) \) and \( \text{As}(3b) \) states \((\text{which differ predominantly in the proportions of} \ s \text{and} \ p \text{character}) \) is reduced, approaching degeneracy at about \( \lambda \approx 0.6 \). Hence, the relatively small \( s, p \) nonlocality in \( \text{As} \) and \( \text{Ga} \) \((\text{Fig. 1}) \) seems to be responsible for the formulation of two stable \( \text{As} \)-derived surface states in this energy region.

Finally, when the \( p, d \) nonlocality is reduced and the \( s \) pseudopotential is scaled, the low \( \text{As}(1) \) and \( \text{As}(2) \) surface states move upwards with \( d \epsilon / d \lambda \approx 0.1 \) eV, but the bulk band at \(-12\) \text{eV} moves rigidly with these surface states. The scale of the \( s \) potential determines therefore the position of the lowest \( s \)-derived surface states with respect to vacuum but the positions relative to the nearest band edges remain nearly invariant under scaling.

I end this paper with a theoretical note on the validity of the density-functional approach\(^ {6} \) used here and in other pseudopotential \([e.g., \text{Ref. 15}] \), as well as all-electron\(^ {16} \) calculations, to describe localized surface states. As can be seen from Eqs. (7)–(9), the density-functional approach includes in the effective screening \( W_{\text{eff}}[n(\mathbf{r})] \) the unphysical Coulomb repulsion of an electron in state \( \psi \) with itself: \( V_{\text{ex}}[\psi] \), as well as the exchange-correlation attraction of an electron with itself: \( V_{\text{ex}}[\psi^2] \). These self-interaction terms are negligible only for diffused orbitals. The partial success of the density-functional formalism in describing localized states rests on the fact that often most of the \textit{positive} self-Coulomb term is canceled by the \textit{negative} self-exchange-correlation terms. In general, however, recent calculations\(^ {18} \) have shown that \( V_{\text{ex}}[\phi] > |V_{\text{ex}}[\phi^2]| \) for a wide range of localized orbitals \( \psi(\mathbf{r}) \), leading therefore to anomalously \textit{high} orbital energies for localized
states in the density-functional approach, relative to diffused itinerant states (which have a vanishing self-interaction).

It has been recently pointed out\(^\text{19}\) that one can go beyond the local-density formalism by defining a new energy functional in which self-interaction effects are canceled self-consistently. This leads to a modified state-dependent screening which replaces Eq. (7):

\[
W_{\text{sc}}[\rho(\mathbf{r}), \psi^\dagger(\mathbf{r})] = V_{\text{el}}[\rho(\mathbf{r})] + V_{\text{ex}}[\rho(\mathbf{r})] \\
+ V_{\text{el}}[\rho(\mathbf{r})] - V_{\text{calc}}[\psi^\dagger(\mathbf{r})] \\
- V_{\text{el}}[\psi^\dagger(\mathbf{r}), \xi = 1] \\
- V_{\text{el}}[\psi^\dagger(\mathbf{r}), \xi = 1],
\]

where \(\xi\) denotes spin polarization. Applications to many atoms have shown that the orbital energies obtained with (19) are 1–3 eV lower than those obtained with Eq. (7), even for the valence electrons. In addition, many of the systematic anomalies characterizing the local-density formalism have been shown to be removed by this self-interaction corrected scheme.

Since some of the surface states obtained for semiconductors have a localization range characteristic of an atomic scale \(\{\text{e.g., states appearing in gaps in the projected band structure such as Ga(1) and Ga(2) in Fig. 2}\}\), one may expect that self-interaction corrections for these states would be a non-negligible fraction of that found for atoms. Indeed, the energy of the Ga(1)-Ga(2) states calculated here is about 0.5–0.7 eV higher than the experimental value.\(^\text{22}\) Hence, whereas many electronic-structure calculations for semiconductor surfaces using a density-functional screening Eq. (7) \(\{\text{e.g., Ref. 49 and references therein}\}\) have produced an overall agreement with experiment (sometimes via semiempirical parametrization of the pseudopotential or scaling the exchange potential), as pointed out first by Schrieffer,\(^\text{20}\) non-negligible corrections can result from the physical mechanisms underlying Eq. (19).

ACKNOWLEDGMENT

I thank J. van Laar and R. Bauer for comments and discussions and J. Chelikowsky for criticism of the manuscript.

\(^34\)F. Herman and S. Skillman, Atomic Structure Cal-
46L. Pauling, Nature of the Chemical Bond (Cornell University, Ithaca, 1960).