Linear combinations of constituent states: a first-principles Luttinger model



1996 J. Phys.: Condens. Matter 8 11029

(http://iopscience.iop.org/0953-8984/8/50/027)

J. Phys.: Condens. Matter **8** (1996) 11029–11039. Printed in the UK

# Linear combinations of constituent states: a first-principles Luttinger model

Sverre Froyen
National Renewable Energy Laboratory, Golden, CO 80401, USA

Received 24 July 1996

**Abstract.** We present a method where Luttinger-like parameters extracted from first-principles calculations on small prototype systems are used to calculate near-band-gap states of much larger semiconductor heterostructures. The method can be used for small- and large-scale heterostructures with equal first-principles accuracy and it illustrates approximations that are implicit in the Luttinger model. GaAs–AlAs(001) superlattices are used to illustrate the method.

## 1. Introduction

Traditionally, the band structure of large semiconductor heterostructures, such as long-period superlattices, has been determined using empirical methods like the Luttinger model [1]. For smaller systems, where the accuracy of the Luttinger model is questionable, first-principles methods such as the linearized augmented plane-wave (LAPW) method [2] and the *ab initio* pseudopotential method [3] are commonly used. Because the latter methods are based upon the local density approximation [4], they give results that differ from the empirical methods, rendering comparisons difficult. It is also hard to ascertain the errors associated with the empirical methods.

In this paper, we describe a formalism that extracts Luttinger-like parameters from first-principles calculations on one or more prototype systems. These parameters can then be used to calculate the large-scale heterostructures with first-principles accuracy. The formalism also provides a connection between first-principles methods and the Luttinger model and helps clarify some of the approximations that are implicit in the Luttinger model.

Another method that is currently being used for calculations on similar systems is the spectrum folding method [5] that solves directly the Schrödinger equation derived from the empirical and semi-empirical pseudopotential methods.

## 2. Formalism

### *2.1. Introduction to basis sets*

For the purposes of this discussion, we consider a superlattice in A and B, where A and B are two semiconducting zincblende (or diamond) constituents. Note, however, that generalizations to more complex three-dimensional structures as well as structures containing additional constituents are straightforward. We also assume, for now, that the two materials are lattice matched, but see subsection 2.7 for a discussion of how to remove this constraint. The method we propose is based on the standard $k$-dot-$p$ theory [1]. In $k$-dot-$p$-like models,

0953-8984/96/5011029+11$19.50  

the superlattice wavefunction $\Psi$ is expanded in the periodic part of the constituent Bloch functions (usually chosen from the $\Gamma$ point of the Brillouin zone) $u_{n,0}$ as

$$\Psi(r) = \sum_n F_n(r) u_{n,0}(r) \tag{1}$$

where $F_n(r)$ are so-called envelope functions, commonly assumed to vary slowly. We assume for now that that basis states $u_{n,0}$ are the same for materials A and B and therefore the same throughout the superlattice. This basis set is commonly referred to as the $k_0$ representation and is also called the Luttinger–Kohn basis. For periodic systems, the envelope functions $F_n(r)$ can be expanded in the plane waves. The result is

$$\Psi(r) = \sum_n \sum_k F_{n,k} u_{n,0}(r)\, \mathrm{e}^{\mathrm{i}kr}. \tag{2}$$

The $k$-vector sum is discrete because the superlattice periodicity makes each $k$-vector in the sum a reciprocal lattice vector of the supercell. The sum is also finite because extending the $k$-vector sum beyond the first Brillouin zone of the bulk (zincblende) is equivalent to adding additional terms to the band sum $n$.

An alternative expansion is to use $u_{n,k}(r)$ in place of $u_{n,0}(r)$. This gives

$$\Psi(r) = \sum_n \sum_k F'_{n,k} u_{n,k}(r)\, \mathrm{e}^{\mathrm{i}kr} \tag{3}$$

and corresponds to the Wannier–Stark basis in $k$-dot-$p$ theory [1]. (3) is an expansion in the full Bloch states of the constituents (ignoring for now differences between A and B). Notice that when the band sum $n$ is extended to all bands, the basis forms a complete set of states (this is true whether $u_{n,0}$ or $u_{n,k}$ is used). By judiciously selecting the basis states, one can therefore obtain an arbitrarily accurate $\Psi$. The challenge is to select the smallest set of basis states able to represent $\Psi$ with a given accuracy, and herein lies the difference between (2) and (3). In general, the $u_{n,k}$ basis will afford the better accuracy. Consider, for instance, the case where materials A and B are identical. Using the $\Gamma$-derived $u_{n,0}$ basis a large number of basis states is required to describe the lowest conduction band at the X point, whereas with the $u_{n,k}$ basis a single state is all that is needed.

### *2.2. The Luttinger model*

Although it is commonly implied that the eight-band Luttinger model uses the $u_{n,0}$ of the $k_0$ representation, in reality it uses a combination of $u_{n,0}$ and $u_{n,k}$. The Luttinger basis states are never specified explicitly, but they are assumed to be eigenstates of the constituent Hamiltonian in both regions A and B. The Hamiltonian is parametrized with parameters taken from experiment. The parameters include $\Gamma$ energy levels, effective masses, and inter-band momentum matrix elements. In the pure $k_0$ representation, however, the effective masses are exactly equal to the free electron masses and the inter-band momentum matrix elements are strictly linear in $k - k_0$ [1]. The correct effective mass for a particular band arises from the inter-band interactions. An eight-band basis, however, is too small to produce good effective masses this way and the Luttinger model instead resorts to explicit effective mass parameters. Formally, this introduces a $k$-dependence in the basis. In the limit of a single-band, the Luttinger-type basis is equivalent to the Wannier–Slater basis but, because of its parametrization, in the effective mass approximation only.

It is unlikely that the two constituents, A and B, are similar enough for their basis states to be identical. Since the Luttinger model parameters are fitted to reproduce the properties of the constituents, it is implicit that its basis states are those of material A in region A and those of material B in region B. The Luttinger model, therefore, not only ignores discontinuities

in the basis functions at the boundary between A and B, but also ignores the fact that the resulting basis is non-orthogonal. The $k$-dependence of the basis further exacerbates these deficiencies and necessitates *ad hoc* remedies such as the commonly used $(1/m)$-discontinuity in the envelope-function derivative at material boundaries. Nevertheless, the Luttinger model is shown to give excellent result for simple structures such as long-period superlattices [6].

### *2.3. The virtual crystal approximation*

Another common basis choice is that of using the states obtained from the virtual crystal approximation (VCA). The formalism is simple. There are no boundary matching problems and the method can easily be combined with first-principles calculations. A large number of bands, however, is required to represent accurately the constituents A and B. For long-period superlattices, any finite-size VCA basis choice will therefore always introduce large relative errors.

### *2.4. The linear combination of constituent states*

We can combine the straightforwardness of the VCA and the constituent accuracy of the Luttinger model by using the states from A as well as the states from B throughout the superlattice. Generalizing (2),

$$\Psi(r) = \sum_n \sum_k F^A_{n,k} u^A_{n,k}(r)\, \mathrm{e}^{\mathrm{i}kr} + \sum_n \sum_k F^B_{n,k} u^B_{n,k}(r)\, \mathrm{e}^{\mathrm{i}kr}. \tag{4}$$

The resulting basis set is, of course, over-complete and must be orthonormalized before being used. By construction, the basis set in (4) is able to describe exactly the constituent A and B. As indicated in subsection 2.1, the Wannier–Slater basis will produce the best accuracy with the smallest number of states.

Inserting our finite basis set into the superlattice Schrödinger equation, $H\Psi = E\Psi$, leads to a secular equation in the unknown energy level $\varepsilon$ and the coefficients $F_{n,k}$,

$$\sum_{n,k} \big[ H\big(n',\, k';\, n,\, k\big) - \varepsilon S\big(n',\, k';\, n,\, k\big) \big] F_{n,k} = 0. \tag{5}$$

In order to simplify the notation, the sum over A and B has been folded into the sum over $n$. We have included an explicit overlap matrix $S$ to handle the case where $u^A_{n,k}(r)$ and $u^B_{n,k}(r)$ are not pre-orthogonalized.

The Hamiltonian matrix elements $H(n',\, k';\, n,\, k)$ can be evaluated constituent unit cell by constituent unit cell. With $R_i$ denoting the position of cell $i$, a matrix element can be written

$$\begin{aligned} H\big(n',\, k';\, n,\, k\big) &= \int \mathrm{d}r\, u^*_{n',k'}(r)\, \mathrm{e}^{\mathrm{i}k'r} H(r) u_{n,k}(r)\, \mathrm{e}^{\mathrm{i}kr} = \sum_i \int_{\Omega_i} \mathrm{d}r\, u^*_{n',k'}(r)\, \mathrm{e}^{-\mathrm{i}k'r} H(r) u_{n,k}(r)\, \mathrm{e}^{\mathrm{i}kr} \\ &= \sum_i \mathrm{e}^{\mathrm{i}(k-k')R_i} \int_{\Omega_i} \mathrm{d}r\, u^*_{n',k'}(r)\, \mathrm{e}^{\mathrm{i}k'r} H\big(R_i + r\big) u_{n,k}(r)\, \mathrm{e}^{\mathrm{i}kr} \\ &= \sum_i \mathrm{e}^{\mathrm{i}(k-k')R_i} H_i\big(n',\, k';\, n,\, k\big). \end{aligned} \tag{6}$$

The sub-matrix elements $H_i(n',\, k';\, n,\, k)$ can be computed either using the potential of the constituent at position $R_i$ and adjusting for the band offsets or, better, using the potential from a prototype short-period superlattice where band offsets as well as interface potentials are naturally included. In section 3 we give an example of this. The evaluation of the

overlap matrix elements is simplified by a lack of cell index dependence. The overlap matrix is therefore diagonal in $k$ and $k'$.

The matrix size in (5) is equal to the number of bands multiplied by the number of $k$-vectors multiplied by the number of constituents. The latter is limited by the number of superlattice reciprocal cell vectors inside the first Brillouin zone of the constituents, which in turn is equal to the number of constituent unit cells in the superlattice cell. For coarse-grained structures such as long-period superlattices, it is frequently possible to truncate the $k$-vector sum to a few $k$-vectors close to the constituent band extrema.

For analysis purposes various eigenvector projections can be defined. The simplest is a projection on the constituent basis states. For instance, the projection of $\Psi(r)$ on the basis state $u_{n',k}(r)$ is

$$P_{n'k'} = \int \mathrm{d}r u^*_{n',k'}(r)\,\mathrm{e}^{-\mathrm{i}k'r}\Psi(r) = \sum_n S\big(n',\ k';\ n,\ k'\big)F^A_{n,k'}. \tag{7}$$

Another useful quantity is the cell projection. This allows us to produce envelope-function-like plots. Projecting on constituent unit cell $i$, we obtain

$$P\big(R_i\big) = \int_{\Omega_i} \mathrm{d}r|\Psi|^2 = \sum_{n',k'}\sum_{n,k} F_{n',k'}S_i\big(n',\ k';\ n,\ k\big)F_{n,k}\,\mathrm{e}^{\mathrm{i}(k-k')R_i} \tag{8}$$

where $S_i$ is defined analogously to $H_i$ in (6). $S_i$ is *not* diagonal in $k$ and $k'$ because the integral $\Omega_i$ is over only a single constituent unit cell.

The envelope functions provided by (8) can be used to model electron or hole densities in degenerate semiconductor heterostructures (where electrons occupy conduction band states or holes occupy valence band states). Poisson's equation, suitably screened and possibly also including exchange and correlation effects, can be solved to determine the resulting (slowly varying) potential. Assuming the potential can be treated as a constant $V_i$ over the constituent unit cell $i$, (6) can be modified by simply adding the term $V_iS_i(n',k';n,k)$ to $H_i(n',k';n,k)$. Since potential changes would modify the eigenstates, this procedure would have to be performed selfconsistently. A one-band effective-mass version of this procedure was recently used by Fong *et al* to model coupled quantum dots [7].

### *2.5. Approximations to the matrix element*

As we shall see in section 3, (5) and (6) produce excellent results (as compared to full calculations) when applied to short-period GaAs–AlAs superlattices. For longer periods, however, the time spent evaluating the Hamiltonian matrix elements becomes excessive. The number of basis states scales linearly with the size of the superlattice, and, since all basis states in general are coupled, the number of matrix elements scales with the square of the system size. For simple large-scale structures, the near-band-gap states are mostly made up of basis states near the band extrema of the constituents, and this fact can be used to reduce the number of basis states. Still, it is useful to have a simple and fast interpolation scheme for evaluating the matrix elements.

While it would be optimal to have a single interpolation formula that works throughout the Brillouin zone, band crossings make such a general scheme difficult to implement. Borrowing instead from the Luttinger model, we choose to expand the matrix elements in a power series of $k - k_0$ around some point $k_0$. Note that we are not limited to a single $k_0$ expansion point. We are free to choose as many expansion points as are necessary to obtain a required accuracy. In the example, below, we expand around Γ, X, and one or two points between Γ and X. This, as we shall see, gives us a good description of all states between Γ

and X. For very long-period superlattices, where we are interested in near-band-gap states only, we expect that Γ and X will suffice.

In our effort to obtain an accurately working power series we made the following observations.

(i) Since we are working with a non-orthogonal over-complete basis set, we need to expand the overlap matrix as well as the Hamiltonian matrix. The overlap matrix contains $n^2$ elements that are calculated from and depend on $n$ states. The overlap matrix elements are therefore not independent. When approximating the overlap matrix, care must be taken so that these $n^2$ overlap matrix elements remain mutually consistent. Otherwise the positive definiteness of the matrix is lost and the result is a catastrophic loss of accuracy. In order to avoid this, we will limit the $k$-dependence of the basis functions to first order in $k - k_0$. The resulting overlap matrix can be evaluated exactly by a second-order expansion.

(ii) Using the $k_0$-representation is attractive for two reasons. First, the overlap matrix is independent of $k$ so the overlap matrix consistency problem described in the previous paragraph is trivially solved. Second, it follows from the $k$-dot-$p$ theory that the Hamiltonian matrix elements are (aside from the effects of the non-local pseudopotentials) exact to second order in $k - k_0$. The disadvantage of the $k_0$-representation is that, because each band in isolation has a mass equal to the free electron mass, a large number of bands is needed to obtain the correct effective masses. Not counting spin degeneracies, fewer than seven bands is useless and 15 bands are needed for an accurate description of the states near the Γ point in GaAs and AlAs. If the number of bands is reduced to, say, four (equivalent to the eight-band Luttinger model), the lack of conduction band p states causes the effective mass of the valence-band heavy-hole states to have the wrong *sign*. They disperse upwards and close the semiconductor gap.

(iii) An alternative which also satisfies (i) is to use an approximate Wannier–Slater type basis with basis states that are correct to first order in $k - k_0$, that is, a basis of the form

$$\tilde{u}_{n,k}(r) = u_{n,0}(r) + u'_{n,0}(r)\big(k - k_0\big). \tag{9}$$

$u'_{n,0}(r)$ is computed from the exact $u_k(r)$ (at $k = k_0 \pm \delta k$) by re-normalizing $u_k(r)$ using the modified norm

$$\big\langle u_k(r) | u_{n,0}(r) \big\rangle = 1 \tag{10}$$

and numerically differentiating using finite differences. The resulting basis $\tilde{u}_{n,k}(r)$, which is neither normalized nor necessarily orthogonal, contributes to the overlap matrix. This choice of norm, (10), makes $u'_{n,0}(r)$ as well as the second-order term orthogonal to $u_{n,0}(r)$. Each band, therefore, has the correct effective mass—even in the absence of inter-band interactions. With the Luttinger–Kohn basis, we noted that the matrix elements were accurately described by the second-order expansion in $k - k_0$. Unfortunately, this is not so for the Wannier–Slater basis. The resulting band structure is highly inaccurate beyond the effective mass region.

Borrowing again from the Luttinger model, we want to combine the best features from the two basis sets. We first observe that $u'_{n,0}(r)$ in (9) can be orthogonalized, not only to $u_{n,0}(r)$ of band $n$, but to all the states $u_{m,0}(r)$ that are included in the basis at $k_0$. After the Hamiltonian has been diagonalized, the resulting wavefunctions are still correct to first order in $k - k_0$. The benefit of this procedure is a decrease in $k$-dependence of $\tilde{u}_{n,k}(r)$ and a corresponding increase in the inter-band interactions. This *hybrid* basis, i.e., (9) with the additional orthogonalization, combines the best properties of the Luttinger–Kohn and the Wannier–Slater bases. For a large number of basis states, it reduces to the Luttinger–Kohn

basis, with the well behaved matrix elements that are accurately described to second order in $k - k_0$. For a smaller number of basis states it approaches the Wannier–Slater basis with its correct description of the effective masses.

Its only disadvantage is that the basis depends on the particular bands that are included in the basis. Like the approximate Wannier–Slater basis in (iii), above, these basis states are not necessarily orthonormal.

### *2.6. Summary*

Consider finally how the basis, (9) with modifications, can be used to describe a large mesoscopic system. First, choose the basis functions. This involves selecting which bands to include and also the $k_0$ expansion points. The choice can later be checked by performing calculations on short-period superlattices. Second, compute the basis functions from the first-principles calculations for the constituents. Third, create the superlattice Hamiltonian. This is best accomplished by building the Hamiltonian unit cell by unit cell either by using potentials from the constituent calculations and adjusting for the band offset, or by using first-principles calculations for a short- to medium-period prototype superlattice. If the latter is used, interface Hamiltonians, as well as band offsets, are obtained naturally. Fourth, use (6) to compute the matrix elements of the basis functions with the Hamiltonian. At this point it may be possible to reduce the number of $k$-vectors. Finally, the secular equation, (5), is diagonalized using either standard methods or possibly an iterative method, such as the variation of the conjugate gradient method used by Wang and Zunger for finding near-band-gap states in their EPM calculations [5].

### *2.7. Other considerations*

For superlattices with lattice-mismatched constituents, one must decide how a basis set belonging to one material is continued into the other material. The simplest choice is to scale the basis set with the lattice parameters, i.e., $u(r)\,\mathrm{e}^{\mathrm{i}kr} \rightarrow cu(\varepsilon r)\,\mathrm{e}^{\mathrm{i}k\varepsilon r}$ where $\varepsilon$ is the strain tensor that transforms the lattice parameters of one material into those of the other material and $c$ is a normalization constant. This scaling can be done on a constituent cell by constituent cell basis that leaves the basis state overlaps invariant. For instance, for a basis set of material A, imagine a supercell that is isomorphic to the actual heterostructure supercell, but that contains only material A and where every constituent unit cell has a size and shape equal to that of material A [8]. Consider an A-type basis in this supercell. Since the material is pure A in all respects, the basis state is perfectly commensurate with the lattice. We now transform this supercell into the actual heterostructure supercell. Do this by 'attaching' the basis set to the constituent unit cell boundaries while transforming the shape and size of every constituent unit cell from an A cell to the appropriate heterostructure cell. Provided the A basis state within each constituent unit cell is scaled uniformly, basis state inner products within each constituent cell do not change and, hence, the overlap matrix is invariant. This cell by cell scaling does, however, create discontinuities in the derivative of the basis state at the material interfaces. Near interfaces, there may also be atomic displacements within the constituent cells. In order to take these interface effects into account, further scaling of the basis states may be necessary. This can be achieved by using a position-dependent strain tensor. 'attaching' the basis states to the atoms while scaling.

Finally, for heterostructures more complicated than one-dimensional superlattices, we must generalize our method to three dimensions. The $k$-vector sums are now three

dimensional but remain bounded by the constituent reciprocal unit cells. The $k$-vector extrapolation of the basis states, (9), and the matrix elements must be generalized to three dimensions. This is straightforward, except when bands are degenerate. Band degeneracies complicate the computation of $k$-vector gradients because a degeneracy usually implies a singularity in $k$. For instance, approaching the valence band maximum at $\Gamma$, the limit of the eigenstates as $k$ approaches zero depends on the direction from which $\Gamma$ is approached. We can express this as

$$\lim_{k\to 0}\big[u_k\big] = U_{\hat{k}}\big[u_0\big] \tag{11}$$

where $[u_0]$ denotes a set of eigenstates that are degenerate at $k = 0$, $[u_0]$ is a reference choice for the degenerate set, and $U_{\hat{k}}$ is a unitary matrix ($3 \times 3$ for the threefold-degenerate valence band maximum at $\Gamma$) that depends on the direction of $k$, that is, $\hat{k}$. For simplicity we choose $\Gamma$ for the expansion point ($k_0 = 0$), but generalizations are obvious. We can use $U_{\hat{k}}$ to simplify the calculation of the $k$-vector gradients. Instead of operating with the states $[u_k]$, use instead the states

$$\big[v_k\big] = U_{\hat{k}}^{-1}\big[u_k\big]. \tag{12}$$

Although these states are no longer eigenstates of the constituent Hamiltonian, they are equivalent to $[u_k]$ when used as basis states. The set $[v_k]$ is furthermore non-singular as $k$ approaches zero and each state can therefore be differentiated separately from the others.

## 3. Examples

For testing the new method we will use the traditional $(\mathrm{GaAs})_N\mathrm{AlAs}_N(001)$ superlattices. The first-principles calculations are performed using semi-relativistic norm-conserving pseudopotentials [3, 9] and the local density approximation [4]. The calculations of the constituents, GaAs and AlAs, as well as the prototype superlattice, $(\mathrm{GaAs})_4(\mathrm{AlAs})_4$, are performed with a conjugate gradient program using a plane-wave kinetic-energy cutoff of 20 Ryd.

From the constituents, we obtain the basis functions, and the prototype superlattice gives us the Hamiltonian needed to compute the sub-matrix elements of (6). We make the assumption that the potential in the constituent unit cell $i$ depends on the composition at site $i$ as well as the sites $i-1$ and $i+1$. For any superlattice, therefore, there are eight types of cell (potentials). With obvious notation, we denote these ${}_A\mathrm{A}_A$, ${}_A\mathrm{A}_B$, ${}_A\mathrm{B}_A$, etc. The $N = 4$ prototype superlattice can be used to calculate the matrix elements of the constituents, ${}_A\mathrm{A}_A$ and ${}_B\mathrm{B}_B$, as well as the various interface matrix elements with the exception of ${}_A\mathrm{B}_A$ and ${}_B\mathrm{A}_B$. Note that band offsets are automatically accounted for.

We will test with two types of basis set. The first and most accurate is the exact Wannier–Slater basis in which every matrix element is calculated separately, i.e., without resort to power-series extrapolation. We use four valence and four conduction bands for a total of eight bands, not counting spin degeneracies. We use 40 $k$-vectors along $\Gamma$–X–$\Gamma$. For this basis set we tabulate all the matrix elements ($H$ and $O$). This basis set allows us to calculate the $\Gamma$ states of any superlattice whose period is 40 or a divisor of 40. As we shall see, this basis set produces very accurate results for short-period superlattices where direct comparisons with first-principles methods are possible. We shall therefore use the results from this basis set as a reference for estimating the accuracy of the (presumably) less accurate second type of basis set.

The second basis set type uses the modified form of (9). We generated two basis sets, one containing eight bands and one with four bands. For $k_0$ expansion points, we use the $\Gamma$

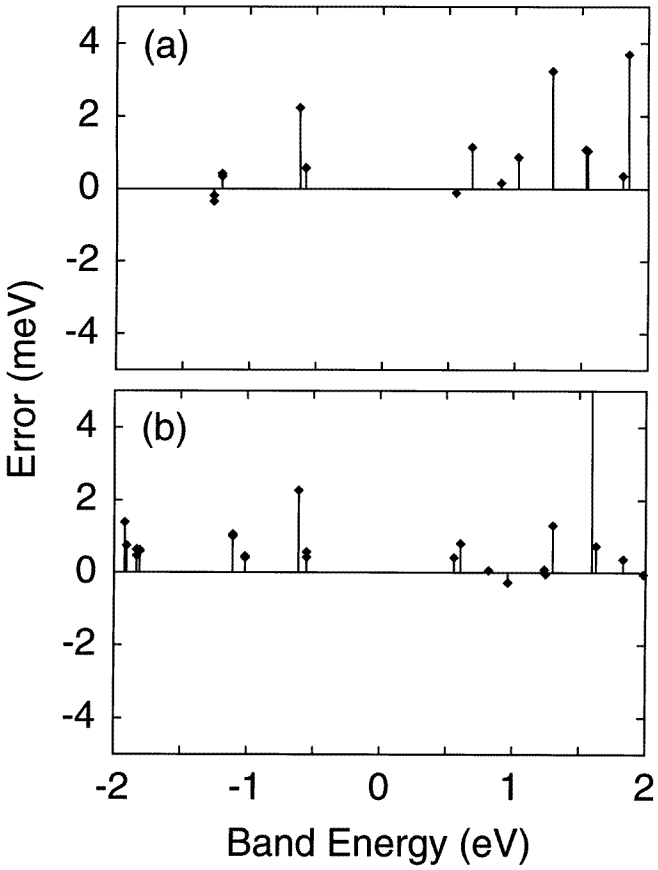


**Figure 1.** Band-energy differences between the first-principles and model calculations as a function of energy for the exact Wannier–Slater basis set. The calculations are performed for (a) $(AlAs)_4(GaAs)_4$ and (b) $(AlAs)_5(GaAs)_5$ superlattices.

and X points plus one or two points between Γ and X. For each constituent and $k_0$ expansion point, the basis-function correction $u'_{n,0}(r)$ in (9) is evaluated by computing three sets of wavefunctions, $u_{n,0}(r)$ and $u_{n,\pm}(r)$, at $k_0$ and $k_0 \pm \delta k$. We used $\delta k = 0.01\,2\pi/a$. $u_{n,\pm}(r)$ is re-normalized using (10), and the difference $(u_{n,+}(r) - u_{n,-}(r))/2\delta k$ is orthogonalized to the other basis states $u_{m,0}(r)$ of the same constituent at $k_0$. The result is $u'_{n,0}(r)$. A table of Hamiltonian matrix elements and their first and second derivatives is subsequently computed from the Hamiltonian of the prototype superlattice. The overlap matrix elements are tabulated as well. For each basis set $k$-vector, the matrix elements are then extrapolated from the closest $k_0$ expansion point.

Starting with the first basis set, in figure 1 we compare the superlattice band energies from $N = 4$ and $N = 5$ model superlattices with those obtained from direct first-principles calculations. The figure shows the band energy difference versus the band energy in a 4 eV region around the band gap. The error is generally less than 1 meV with a couple of exceptions. An error of this magnitude is typical throughout the valence band, whereas for the upper conduction bands the error increases because the basis set was limited to eight bands. For comparison purposes, we have calculated the average error and its standard

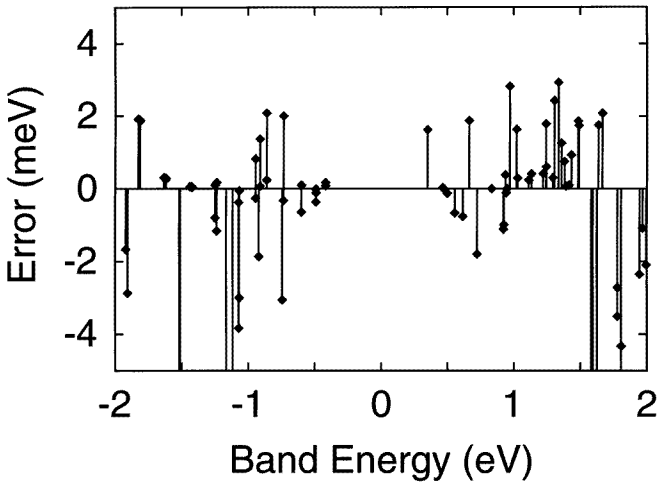


**Figure 2.** Band-energy differences between model calculations using the exact Wannier–Slater basis set and the hybrid basis set with power-series expansion of the matrix elements. The calculation is performed for an $(AlAs)_{20}(GaAs)_{20}$ superlattice.

deviation in a couple of intervals around the gap. These are tabulated in table 1. For longer-period superlattices, where interfaces become less important, we expect the errors to be smaller. Note that the basis, by construction, is able to describe exactly both of the constituents.

**Table 1.** Error when using the Wannier–Slater basis with exact matrix elements and eight bands for $(AlAs)_n(GaAs)_n$ superlattices. The error is given in millielectron volts as the mean of the error and the standard deviation of the error in the interval indicated (measured from approximately midgap). An overall shift has been applied to the energy levels so that the mean error for the largest interval is zero.

| | Error in interval | | |
|---|---|---|---|
| $n$ | $\pm 0.6$ eV | $\pm 1.2$ eV | $\pm 1.8$ eV |
| 4 | $-0.5 \pm 0.3$ | $0.0 \pm 0.7$ | $0.0 \pm 0.9$ |
| 5 | $-0.4 \pm 0.1$ | $-0.2 \pm 0.6$ | $0.0 \pm 1.2$ |

For the second basis type, which uses the hybrid basis with power-series expansion of the matrix elements, the error is in general larger than for the first type. We will use the results from the first basis type for comparison. Compared to a first-principles calculation, there is therefore an uncertainty of the order of 1 meV in the error estimates. In figure 2, as an example, we show such a comparison for an $N = 20$ superlattice. The basis that was used contains eight bands and all (40 for $N = 20$) $k$-vectors. We used six $k_0$ expansion points around which the matrix elements are expanded: Γ, (1/3)X, (2/3)X, X, (4/3)X, and (5/3)X. We show only a 4 eV interval around the band gap but the errors in this interval are typical of a much wider range, including the whole of the valence band.

If we are interested only in states very close to the band gap, the basis size can be reduced. This is illustrated in table 2, which summarizes the errors in various-sized intervals

**Table 2.** Error when using the hybrid basis type, truncating the number of bands in the basis and the number of $k$ expansion points, for $(AlAs)_n(GaAs)_n$ superlattices. The error is given in millielectron volts as the mean of the error and the standard deviation of the error in the interval indicated (measured from approximately midgap).

| | | | Error in interval | | |
|---|---|---|---|---|---|
| $n$ | No of bands | No of $k$-points | ±0.6 eV | ±1.2 eV | ±1.8 eV |
| 4 | 8 | 6 | $0.0 \pm 0.1$ | $0.3 \pm 0.7$ | $-1.9 \pm 5.1$ |
| | 8 | 4 | $2.7 \pm 7.9$ | $-10 \pm 16$ | $-13 \pm 21$ |
| | 4 | 6 | $2.0 \pm 0.1$ | $-3.0 \pm 6.4$ | $1.9 \pm 7.6$ |
| 5 | 8 | 6 | $0.1 \pm 0.5$ | $0.2 \pm 0.8$ | $-1.8 \pm 5.0$ |
| | 8 | 4 | $0.0 \pm 0.3$ | $-0.7 \pm 1.1$ | $-0.3 \pm 2.0$ |
| | 4 | 6 | $2.0 \pm 2.4$ | $4.5 \pm 8.1$ | $-0.4 \pm 12$ |
| 20 | 8 | 6 | $0.0 \pm 0.6$ | $-0.5 \pm 2.2$ | $-1.0 \pm 3.5$ |
| | 8 | 4 | $0.0 \pm 0.6$ | $-3.8 \pm 10$ | $-6.5 \pm 17$ |
| | 4 | 6 | $-0.6 \pm 1.0$ | $-0.1 \pm 2.4$ | $-0.2 \pm 3.0$ |

around the gap. In general the results are more accurate near the band gap. This is to be expected since those states are mostly derived from constituent states near Γ and (for odd $n$) X, which are always included in the $k_0$ expansion point set.

We have truncated the basis in two ways: first by reducing the number of $k_0$ expansion points and second by reducing the number of bands. The $k_0$ expansion point set labelled '4' corresponds to Γ, (1/2)X, X, and (3/2)X. Note that it is only the $k$-points used in the power-series expansion that have been changed. The $k$-point set used for the eigenvectors and the Hamiltonian remains the same. The number of bands is reduced from eight to four by including only the top two valence bands and the bottom two conduction bands except at the Γ point, where three valence bands plus one conduction band were used. We observe that reducing the number of bands is less damaging than reducing the number of $k_0$ expansion points. Note the striking difference in accuracy between $N = 4$ and $N = 5$ when four expansion points are being used. The $N = 4$ superlattice needs basis states at multiples of (1/8)X, which with four expansion points fall exactly in the middle between the expansion points. Increasing the superlattice period to five or increasing the number of expansion points reduces the extrapolation range and improves the result dramatically.

When using the reduced set of bands, we sometimes observe spurious 'ghost' states. The number of such states can be reduced by using a singular-value decomposition of the overlap matrix and then eliminating linear combinations of basis states that have small eigenvalues. What is small? Our tests show that $10^{-4}$ is a good compromise that eliminates ghost states without appreciably increasing the overall error. A few ghost states may still occur when the number of bands in the basis set is decreased. These states can be identified by their small projection (7) on either of the two constituent bases and they can thereby be eliminated.

Table 3 compares the accuracy of the basis choices discussed in subsection 2.5. It illustrates the assertions that were made in that section. The hybrid basis, used above, is clearly superior to the other choices. As shown in subsection 2.5, the $k$-independent Luttinger–Kohn basis completely ignores the interaction with states outside the basis set. We see that for this basis eight basis states per constituent is not sufficient to describe accurately the states around the band gap. A further reduction in the number of basis states produces completely unacceptable results. The Wannier–Slater basis also produces poor

**Table 3.** Error as a function of basis for $(AlAs)_{20}(GaAs)_{20}$ superlattices. For all basis types we use eight bands and six expansion $k$-points. The error is given in millielectron volts as the mean of the error and the standard deviation of the error in the interval indicated (measured from approximately midgap).

| Basis type | Error in interval | | |
|---|---|---|---|
| | ±0.6 eV | ±1.2 eV | ±1.8 eV |
| Hybrid | $0.0 \pm 0.6$ | $-0.5 \pm 2.2$ | $-1.0 \pm 3.5$ |
| Luttinger–Kohn | $3.8 \pm 2.7$ | $10 \pm 11$ | $16 \pm 18$ |
| Wannier–Slater | $-11 \pm 21$ | $13 \pm 44$ | $26 \pm 61$ |
| GaAs only | $24 \pm 30$ | $34 \pm 25$ | $47 \pm 36$ |
| AlAs only | $90 \pm 34$ | $71 \pm 37$ | $67 \pm 35$ |

results because of a strong non-parabolicity in its matrix-element $k$-dependence. Finally, we show that neither a pure GaAs basis nor a pure AlAs alone can produce satisfactory results.

## 4. Summary

We have presented a new method whereby first-principles calculations on small prototype semiconducting systems can be expanded to much larger systems. Luttinger-like parameters extracted from the small first-principles calculations are used to assemble Hamiltonian matrices for much larger heterostructures. Diagonalization of the resulting matrix produces energy levels and envelope-type wavefunctions for near-band-gap states of the heterostructure.

## Acknowledgments

The author thanks L-W Wang, S Zhang, and A Zunger for useful discussions and suggestions. The author also acknowledges helpful discussions with K Kunc and support by a grant from NATO. The work was supported by US DOE under contract No DE-AC36-83-CH10093.

## References

[1] Burt M G 1992 *J. Phys.: Condens. Matter* **4** 6651 and references therein
[2] Wei S-H and Krakauer H 1985 *Phys. Rev. Lett.* **55** 1200 and references therein
[3] Ihm J, Zunger A and Cohen M L 1979 *J. Phys. C: Solid State Phys.* **12** 4409
[4] Kohn W and Sham L J 1965 *Phys. Rev.* **140** A1133
[5] Wang L-W and Zunger A 1994 *J. Chem. Phys.* **100** 2394
[6] Wood D M and Zunger A *Phys. Rev.* B **53** 7949
[7] Fong C Y, Nelson J S, Hemstreet L A, Gallup R F, Chang L L and Esaki L 1992 *Phys. Rev.* B **46** 9538
[8] The model we have in mind is a coherent tetrahedral semiconductor heterostructure. For other systems there may be no simple way to map material A onto material B.
[9] Kerker G 1980 *J. Phys. C: Solid State Phys.* **13** L189