Elimination of spurious solutions from eight-band \( \mathbf{k} \cdot \mathbf{p} \) theory

Bradley A. Foreman *

Department of Physics, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China

(Received 7 July 1997)

A method is developed for eliminating spurious solutions from eight-band \( \mathbf{k} \cdot \mathbf{p} \) theory. It reduces the bulk dispersion to a cubic equation (quadratic along \( \langle 001 \rangle \) and \( \langle 111 \rangle \)), yet gives results virtually indistinguishable from ordinary \( \mathbf{k} \cdot \mathbf{p} \) theory. A unique operator ordering is established for heterostructures, and example calculations on superlattices show good results. [S0163-1829(97)51244-6]

Multiband \( \mathbf{k} \cdot \mathbf{p} \) models are frequently used to describe the electronic band structure of semiconductors in situations where nonparabolicity is important. It was noted early in the development of envelope structure theory that such models can produce spurious solutions with very large wave vectors. If these solutions are evanescent they are little more than a numerical nuisance, wreaking havoc in computer calculations, but having no physical significance. Spurious oscillatory modes are more troublesome, however, since the resulting model contradicts the most basic of experimental facts: that semiconductors possess a band gap.

The problem is an old one, but no fully satisfactory solution exists, despite a recent resurgence of activity in this area. Three general approaches have been suggested. The first is to modify the Hamiltonian by discarding those terms responsible for the spurious solutions. This costs some accuracy in the band structure, since it is no longer possible to fit all experimental effective masses. A second possibility is to keep the original Hamiltonian (which is accurate near \( k=0 \)), but reject the large-\( k \) solutions as unphysical. This is problematic in heterostructures, since the discarded solutions are needed to satisfy the boundary conditions associated with the model, and it is unclear which boundary conditions should be eliminated for mathematical consistency. The third approach therefore advocates retaining all solutions on the grounds that spurious bands have negligible influence on the properties of bound-state eigenfunctions.

This paper presents an eight-band \( \mathbf{k} \cdot \mathbf{p} \) Hamiltonian for electrons and light holes in one dimension:

\[
H_2 = \begin{bmatrix}
E_c + A_c k^2 & iP k \\
-iP k & E_v + A_v k^2
\end{bmatrix}.
\]

Here \( E_c \) and \( E_v \) are the CB and valence-band (VB) energies, \( P \) is the interband momentum matrix element, and \( A_c \) and \( A_v \) describe free-electron and remote-band contributions to the effective masses. The latter are responsible for the existence of spurious solutions through a term \( A_c A_v k^4 \) in the secular equation. Such solutions (oscillatory when \( A_c A_v > 0 \)) can therefore be eliminated by setting \( A_c = A_v = 0 \), which is the basis of the first approach described above. However, this leaves only one parameter \( P \) to fit both the CB and VB masses.

A better fit is obtained if one sets either \( A_c = 0 \) or \( A_v = 0 \), but not both. This provides two fitting parameters, while still ensuring the existence of a band gap. In three dimensions it is preferable to set \( A_c = 0 \), because the VB mass is anisotropic.

This simple idea may now be applied to a more realistic eight-band Hamiltonian. We begin with a review of standard \( \mathbf{k} \cdot \mathbf{p} \) theory. The basis used here is

\[
[S \uparrow \bigg| X \uparrow \bigg| Y \uparrow \bigg| Z \uparrow \bigg| S \downarrow \bigg| X \downarrow \bigg| Y \downarrow \bigg| Z \downarrow ]^T,
\]

in which the \( \mathbf{k} \cdot \mathbf{p} \) Hamiltonian for bulk crystals is

\[
H_S = \begin{bmatrix} H_A & 0 \\ 0 & H_A \end{bmatrix} + H_{SO},
\]

where \( H_A \) is the 4×4 block

\[
H_A = \begin{bmatrix} H_{cc} & H_{cv} \\ H_{vc} & H_{vv} \end{bmatrix},
\]

and \( H_{SO} \) is the spin-orbit Hamiltonian given in Eq. (4b) of Ref. 15. In (4), \( H_{cc} \) is the scalar

\[
H_{cc} = E_c + A_c k^2,
\]

where \( A_c \) is calculated from the following expression:

\[
A_c = \frac{h^2}{2m_e} - 2P^2/3E_c - P^2/3(E_e + \Delta).
\]

This paper presents an eight-band \( \mathbf{k} \cdot \mathbf{p} \) Hamiltonian for electrons and light holes in one dimension:
Here $m_c$ is the CB mass, $E_g$ is the band gap, $\Delta$ is the spin-orbit splitting, $P = (\hbar/m)(S|p_x|X)$ is the momentum matrix element (known from the CB $g$ factor\textsuperscript{16} or the mass of the split-off band\textsuperscript{7}), and $m$ is the free-electron mass. Returning to (4), $H_{cv} = H_{uv}^c$ is the $1 \times 3$ matrix

$$H_{cv} = [i P k_x \quad i P k_y \quad i P k_z],$$

(7)

where Kanè's $B$ parameter\textsuperscript{1} is neglected. $H_{uv}$ is a $3 \times 3$ matrix with typical diagonal and off-diagonal elements

$$H_{XX} = E_v - \frac{1}{2} \Delta + L' k_x^2 + M(k_y^2 + k_z^2),$$

$$H_{XY} = N' k_x k_y.$$

(8)

The other elements are given by cyclic permutations of $x$, $y$, and $z$. Here $L'$, $M$, and $N'$ are determined from the measured Luttinger parameters $\gamma'_1$, $\gamma'_2$, and $\gamma'_3$ by first finding the modified parameters

$$\gamma_1 = \gamma'_1 - E_v / 3 E_g,$$

$$\gamma_2 = \gamma'_2 - E_v / 6 E_g,$$

$$\gamma_3 = \gamma'_3 - E_v / 6 E_g,$$

(9)

(where $E_v = 2mP^2/\hbar^2$) and then calculating

$$L' = -(h^2/2m)(\gamma_1 + 4 \gamma_2),$$

$$M = -(h^2/2m)(\gamma_1 - 2 \gamma_2),$$

$$N' = -(h^2/2m)(6 \gamma_3).$$

(10)

This Hamiltonian was developed for bulk crystals, but with some modifications\textsuperscript{17} it may be used in heterostructures as well. Assuming that the zone-center Bloch functions are the same in all media (which is convenient but not always justified\textsuperscript{17,18}), the main problem is to determine the correct order of the operator $k = -i \nabla$ with respect to the effective-mass parameters. The customary technique of “symmetrizing” each matrix element of the Hamiltonian\textsuperscript{7} (i.e., requiring $H_{ij} = H_{ji}$) is now known to be incorrect,\textsuperscript{11,17,19} because although the Hamiltonian as a whole is always Hermitian ($H_{ij} = H_{ji}$), its individual elements generally are not. This idea is not new; it was first discussed by Luttinger.\textsuperscript{20}

A systematic derivation\textsuperscript{1} of the operator ordering for heterostructures gives the following results.\textsuperscript{19,20} The diagonal terms involving $A_c$, $L'$, and $M$ are to be treated according to the usual prescription\textsuperscript{21} $Mk_x^2 \rightarrow k_x M k_x$, but the off-diagonal terms involving $N'$ require more care:

$$H_{XY} \rightarrow k_x N' k_y + k_y N' k_x.$$  

(11)

Here $N'_i$ is the contribution to $N'$ from $\Gamma_i$ and $\Gamma_{12}$ bands, while $N_c$ is that from $\Gamma_{15}$ and $\Gamma_{25}$ bands.\textsuperscript{21} The reason for the different ordering of these terms is easy to see from their definitions\textsuperscript{3,17} in terms of momentum matrix elements. The contribution to $H_{XY}$ from a remote band $j$ of $\Gamma_1$ or $\Gamma_{12}$ symmetry is of the form

$$k_x (X|p_y|j)(E - E_j)^{-1}(j|p_y|Y) k_y,$$

(12)

while that from $\Gamma_{15}$ and $\Gamma_{25}$ bands is the opposite:

$$k_y (X|p_x|j)(E - E_j)^{-1}(j|p_x|Y) k_x.$$

(13)

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|c|}
\hline
$E_p$ (eV) & InSb & InAs & InP & GaSb & GaAs \\
\hline
Reference 16 & 24.4 & 22.2 & 20.7 & 27.9 & 28.9 \\
Equation (14) & 23.0 & 21.7 & 18.2 & 23.6 & 24.3 \\
\hline
\end{tabular}
\caption{Comparison of experimental data for $E_p$ (from Ref. 16) with values calculated from Eq. (14). Calculations use the same material parameters as Ref. 16.}
\end{table}

For $P$ is allowed to vary, then the ordering of $H_{cv}$ is no longer trivial. The best choice in this case is to let $H_{cv}$ have precisely the asymmetric form shown in Eq. (7). The reason is that this is the only ordering that yields a mathematically self-consistent Hamiltonian in the limit $A_c \rightarrow 0$. When $A_c = 0$, the spurious solutions disappear,\textsuperscript{22} so if $H_8$ is to remain well-defined, the number of boundary conditions should decrease accordingly. For any ordering other than (7), however, there remains an extra condition that cannot be satisfied when $A_c = 0$ (as shown below). Such singular behavior must be rejected on physical grounds, because clearly $A_c = 0$ does not correspond to any physical singularity in the energy bands. Since (7) is the only ordering that does not require the existence of spurious solutions, it should also tend to minimize their impact even when $A_c \neq 0$.

In contrast with (11), the ordering (7) of $H_{cv}$ was not taken from a first-principles derivation [see Eq. (6.3) of Ref. 17 and Eq. (4.15) of Ref. 18 for two examples]. This is because such derivations account for neither spurious solutions nor the ensuing singularities that can occur when $A_c = 0$. The operator (7) is designed expressly for these problems, so it is more useful in practice.

Having established a Hamiltonian that is valid for all values of $A_c$, we can now implement the technique described earlier—namely, to eliminate spurious solutions by setting $A_c = 0$ and using $P$ to fit $m_c$. With this choice, Eq. (6) gives

$$E_p = \frac{3m/m_c}{2E_g + 1/(E_g + \Delta)}.$$  

(14)

This definition of $E_P$ implicitly assumes that remote-band contributions to $m_c$ exactly cancel the free-electron mass. This is not a bad approximation, and is certainly better than neglecting remote bands entirely. Table I compares Eq. (14) with experimental data\textsuperscript{16} for $E_p$. The worst case is GaAs, which differs by 16%. This change is small, and since $m_c$ is held constant, the band structure for $A_c = 0$ is almost identical to that for $A_c \neq 0$.
RAPID COMMUNICATIONS

H which holds for both the ↑ and ↓ components. One can therefore eliminate \( F_S \) in favor of \( F_X, F_Y, \) and \( F_Z \).

This reduces the eigenvalue equation to the form \( H_6 \mathbf{F} = E \mathbf{F} \), in which the energy-dependent Hamiltonian \( H_6 \) is obtained from \( H_8 \) by deleting the CB rows and columns and replacing \( L' \) and \( N'_+ \) with [cf. (12)]

\[
L(E) = L' + P^2(E - E_c),
\]

\[
N_+(E) = N'_+ + P^2(E - E_c).
\]

\( M \) and \( N_- \) are unchanged. For bulk crystals the operator ordering (11) is irrelevant and one can use \( N(E) = N'_+(E) + N_-(E) \). Although \( F_S \) has been formally eliminated from the eigenvalue equation, it must be included in any calculation involving envelope functions. (For example, the VB envelopes are neither orthogonal nor properly normalized by themselves.)

\( H_6 \) is just a rearrangement of \( H_8 \), so despite its outward appearance as merely a VB Hamiltonian, its solutions are exactly those of \( H_8 \). The only restriction on \( H_6 \) is that \( E \neq E_c \). This is no limitation in practice, because \( E = E_c \) is only a single point, and the solutions of \( H_6 \) tend to a well-defined limit as \( E \to E_c \). In other words, the singularity is a removable one.

In bulk media \( H_6 \) is the same as the VB Hamiltonian of Eppenga, Schuurmans, and Colak.\(^7\) However, their operator ordering for heterostructures is based on the symmetrization postulate \( N = \frac{1}{2} \) rather than direct derivation from \( H_8 \). This postulate leads to significant errors in the VB\(^\text{11,19} \) and breaks down completely near \( E = E_c \). For \( H_6 \) it is better to approximate \( N \) as totally asymmetric (i.e., \( N_+ = N \) and \( N_- = 0 \)). Of course, \( H_6 \) is less sensitive because it treats CB-VB interactions directly; symmetrizing \( N' \) thus has little effect.\(^5,11\)

If eigenvectors are needed, it is easiest to solve \( H_6 \) by transforming it into a linear eigenvalue problem for \( k(E) \).\(^4\) If not, one can use the following analytical solution for the bulk dispersion. To derive it, one diagonalizes \( H_6 \), then orients the spin along an eigenvector of \( H_{66} \) and solves \( H_6 \). The result is a pair of identical (Kramers degenerate) cubic equations to be solved for \( k^2(E) \):

\[
c_3 k^6 + c_2 k^4 + c_1 k^2 + c_0 = 0,
\]

where

\[
c_3 = LM^2 + M[(L-M)^2 - N^2],
\]

\[
c_2 = -\left( \epsilon + \frac{1}{2} \Delta \right) \{ M(2L + M) + [(L-M)^2 - N^2] \},
\]

\[
c_1 = \epsilon (\epsilon + \frac{1}{2} \Delta)(L+2M),
\]

\[
c_0 = -\epsilon^2 (\epsilon + \Delta),
\]

\( \epsilon = E - E_c \), \( \Theta = (k_x^2 k_y^2 + k_y^2 k_z^2 + k_z^2 k_x^2) / k^4 \), and \( \Omega = k_x^2 k_y^2 k_z^2 / k^6 \).

Along \( \langle 001 \rangle \) and \( \langle 111 \rangle \) the cubic equation separates into a heavy-hole band \( \epsilon = A_H k^2 \) and a quadratic equation for the split-off and light-hole/conduction bands:

\[
(A_H A_L) k^4 - [\epsilon (A_H + A_L) + \Delta (A_H + 2A_L)] k^2 + \epsilon (\epsilon + \Delta) = 0.
\]

For \( \langle 001 \rangle \), \( A_H = M \) and \( A_L = L \), while for \( \langle 111 \rangle \), \( A_H = \frac{1}{2} (L + 2M - N) \) and \( A_L = \frac{1}{2} (L + 2M + 2N) \).

These solutions are compared with the \( A_\epsilon \neq 0 \) solutions of \( H_8 \) in Fig. 1. Except for the spurious bands, no difference between the models is visible at this scale.

Boundary conditions at an abrupt junction may be obtained directly from the envelope-function equations.\(^17\) The results are independent of \( \Delta \) and do not couple the spin-up and spin-down components of \( \mathbf{F} \). As an example, consider a \( \langle 001 \rangle \) junction. If \( A_\epsilon \neq 0 \), then \( \mathbf{F} \) and the following functions (the same for each spin) are required to be continuous:

\[
\begin{align*}
& [A_\epsilon \partial_z, 0, 0, 0, 0, 0, i N'_- k_x, F_Y] \\
& [0, 0, M \partial_z, 0, i N'_- k_x, i N_- k_y, F_Y] \\
& [P, i N'_+ k_x, i N'_+ k_y, L' \partial_z, F_Z] \\
\end{align*}
\]

However, if \( A_\epsilon = 0 \), \( F_S \) is no longer continuous.\(^17,23\) One can then use (15) to eliminate \( F_S \) from (20) and replace (20) with continuity of

\[
\begin{align*}
& [M \partial_z, 0, i N_- k_x, F_X] \\
& [0, M \partial_z, i N_- k_y, F_Y] \\
& [i N'_+ k_x, i N'_+ k_y, L' \partial_z, F_Z] \\
\end{align*}
\]
function equations were solved in a plane-wave basis with \(|k|\geq 2.14 \text{ Å}^{-1}\). The results for \(A_s=0\) and \(A_s \neq 0\) agree quite well, except where spurious bands interfere with the \(A_s \neq 0\) solutions. For \(A_s \neq 0\) and \(P = \text{const}\), there are two spurious bands near the CB edge and none in the VB. The "true" CB ground state also has a large spurious component. For \(A_s \neq 0\) and \(P \neq \text{const}\), the spurious CB pair moves higher and mixes strongly with the upper subbands, while two new spurious modes appear in the VB. Not all of the discrepancies between \(A_s=0\) and \(A_s \neq 0\) are due to spurious bands, since some slight differences remain even when the plane-wave basis is cut off at \(|k| = 0.27 \text{ Å}^{-1}\). However, these differences are no more than those between the two \(A_s \neq 0\) models (\(P = \text{const}\) and \(P \neq \text{const}\)). The available evidence thus clearly favors the \(A_s=0\) model.

The main limitation of the \(A_s=0\) model is approximation (14) for \(E_p\). If this fails (e.g., if remote-band effects are strong, or an experiment is sensitive to the precise value of \(E_p\)), then an eight-band \(A_s \neq 0\) model is not likely to be of much use either. In such cases, one should move to a 14-band model.\(^{24,25}\) The present approach, suitably modified, may prove helpful in that model also.

I am grateful to Mike Burt for providing me with Ref. 13 prior to publication. Initial work on this paper was done while the author was at the University of Essex, and was supported by the North Atlantic Treaty Organization under a grant awarded in 1995. Subsequent work at HKUST was supported by RGC Grant No. DAG96/97.SC38.

---

8Electronic address: phbaf@usthk.ust.hk
13M. G. Burt, *Superlattices Microstruct.* (to be published).
14No parameter is included to fit the mass of the spin-orbit split-off band, but the approximation used here is highly accurate.
20J. M. Luttinger, Phys. Rev. 102, 1030 (1956).
21In Kane’s notation,\(^1\) \(N’_s = F’ - G\) and \(N_s = H_1 - H_2\); in the notation of Ref. 19, \(\gamma_{3s} = \sigma - \delta\) and \(\gamma_{3s} = \pi\).
22In general, the spurious solutions of \(H_s\) are oscillatory when \(A_s [L’ + 2(M’ + N’ - L’)] \Theta > 0\) where \(\Theta\) is given below (18)\(^{1}\) and evanescent when this product is negative.
23The rapidly varying spurious solutions generate an effective discontinuity even when \(A_s \neq 0\). Setting \(A_s = 0\) merely makes this effective discontinuity explicit.