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.^{1,2} It was noted early in the development of envelope structure theory³ that such models can produce spurious solutions with very large wave vectors. If these solutions are evanescent³ they are little more than a numerical nuisance, wreaking havoc in computer calculations^{4,5} but having no physical significance. Spurious oscillatory modes^{6,7} 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.²⁻¹³ Three general approaches have been suggested. The first is to modify the Hamiltonian by discarding those terms responsible for the spurious solutions.²⁻⁴ 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.⁸⁻¹² 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.^{5,13} This is better justified¹³ than most people think,^{8,12} but it still runs into trouble with oscillatory modes (as shown below).

This paper presents an eight-band $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian (for the Γ point of zinc-blende crystals) that has no spurious solutions, yet permits an accurate fit to all experimental effective masses.¹⁴ The basic idea is very simple: to set the coefficient of the k^2 term in the conduction-band (CB) matrix element to zero, and fit the CB mass using a spatially varying momentum matrix element. This yields a complex band structure virtually identical with that of ordinary $\mathbf{k}\cdot\mathbf{p}$ theory. However, the dispersion is simpler, being given by a cubic equation for general $\mathbf{k}$ and a quadratic along $\langle 001 \rangle$ and $\langle 111 \rangle$.

For heterostructures, one must also determine the proper arrangement of differential operators with respect to material parameters. The main new feature of this theory is its position-dependent momentum matrix. A unique operator ordering is derived for this case, giving a Hamiltonian that is

Hermitian but not symmetric. This model is used to calculate the band structure of a superlattice, yielding highly satisfactory results.

A good starting point is the following two-band Hamiltonian for electrons and light holes in one dimension:^{3,8,11}

$$H_2 = \begin{bmatrix} E_c + A_c k^2 & iPk \\ -iPk & E_v + A_v k^2 \end{bmatrix}. \quad (1)$$

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.¹ The latter are responsible for the existence of spurious solutions through a term $A_c A_v k^4$ in the secular equation.^{8,11} 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.²⁻⁴ 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.¹ The basis used here is

$$[S\uparrow \ X\uparrow \ Y\uparrow \ Z\uparrow \ S\downarrow \ X\downarrow \ Y\downarrow \ Z\downarrow]^T, \quad (2)$$

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

$$H_8 = \begin{bmatrix} H_4 & 0 \\ 0 & H_4 \end{bmatrix} + H_{so}, \quad (3)$$

where H_4 is the 4×4 block

$$H_4 = \begin{bmatrix} H_{cc} & H_{cv} \\ H_{vc} & H_{vv} \end{bmatrix}, \quad (4)$$

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

$$H_{cc} = E_c + A_c k^2, \quad (5)$$

where A_c is calculated from the following expression:

$$A_c = \hbar^2/2m_c - 2P^2/3E_g - P^2/3(E_g + \Delta). \quad (6)$$

Here m_c is the CB mass, E_g is the band gap, Δ is the spin-orbit splitting, $P = (\hbar/im)\langle S|p_x|X\rangle$ is the momentum matrix element (known from the CB g factor¹⁶ or the mass of the split-off band⁷), and m is the free-electron mass. Returning to (4), $H_{cv} (= H_{vc}^\dagger)$ is the 1×3 matrix

$$H_{cv} = [iPk_x \quad iPk_y \quad iPk_z], \quad (7)$$

where Kane's B parameter¹ is neglected. H_{vv} is a 3×3 matrix with typical diagonal and off-diagonal elements

$$\begin{aligned} H_{xx} &= E_v - \frac{1}{3}\Delta + L'k_x^2 + M(k_y^2 + k_z^2), \\ H_{xy} &= N'k_xk_y. \end{aligned} \quad (8)$$

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 γ_1^L , γ_2^L , and γ_3^L by first finding the modified parameters

$$\begin{aligned} \gamma_1 &= \gamma_1^L - E_p/3E_g, \\ \gamma_2 &= \gamma_2^L - E_p/6E_g, \\ \gamma_3 &= \gamma_3^L - E_p/6E_g, \end{aligned} \quad (9)$$

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

$$\begin{aligned} L' &= -(\hbar^2/2m)(\gamma_1 + 4\gamma_2), \\ M &= -(\hbar^2/2m)(\gamma_1 - 2\gamma_2), \\ N' &= -(\hbar^2/2m)(6\gamma_3). \end{aligned} \quad (10)$$

This Hamiltonian was developed for bulk crystals, but with some modifications¹⁷ it may be used in heterostructures as well. Assuming that the zone-center Bloch functions are the same in all media (which is convenient² but not always justified^{17,18}), the main problem is to determine the correct order of the operator $\mathbf{k} = -i\nabla$ with respect to the effective-mass parameters. The customary technique of "symmetrizing" each matrix element of the Hamiltonian⁷ (i.e., requiring $H_{ij} = H_{ji}^\dagger$) is now known to be incorrect,^{11,17,19} because although the Hamiltonian as a whole is always Hermitian ($H_{ij} = H_{ji}^\dagger$), its individual elements generally are not. This idea is not new; it was first discussed by Luttinger.²⁰

A systematic derivation¹⁷ of the operator ordering for heterostructures gives the following results.^{19,20} The diagonal terms involving A_c , L' , and M are to be treated according to the usual prescription⁷ $Mk_x^2 \rightarrow k_x Mk_x$, but the off-diagonal terms involving N' require more care:

$$H_{XY} \rightarrow k_x N'_+ k_y + k_y N_- k_x. \quad (11)$$

Here N'_+ is the contribution to N' from Γ_1 and Γ_{12} bands, while N_- is that from Γ_{15} and Γ_{25} bands.²¹ The reason for the different ordering of these terms is easy to see from their definitions^{1,17} in terms of momentum matrix elements. The contribution to H_{XY} from a remote band j of Γ_1 or Γ_{12} symmetry is of the form

$$k_x \langle X|p_x|j \rangle (E - E_j)^{-1} \langle j|p_y|Y \rangle k_y, \quad (12)$$

while that from Γ_{15} and Γ_{25} bands is the opposite:

TABLE I. 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.

E_p (eV)	InSb	InAs	InP	GaSb	GaAs
Reference 16	24.4	22.2	20.7	27.9	28.9
Equation (14)	23.0	21.7	18.2	23.6	24.3

$$k_y \langle X|p_y|j \rangle (E - E_j)^{-1} \langle j|p_x|Y \rangle k_x. \quad (13)$$

N'_+ and N_- may be estimated from bulk measurements if one neglects Γ_{25} bands;¹⁹ in that case, $N_- = M - \hbar^2/2m$ and $N'_+ = N' - N_-$.

Last to be considered is H_{cv} . If the zone-center Bloch functions are truly material independent, P should be constant, but the empirical value of P does vary somewhat even amongst the III-V compounds.¹⁶ Most practical implementations of $\mathbf{k} \cdot \mathbf{p}$ theory get around this by choosing one value of P for the entire heterostructure, and then adjusting A_c to give the right effective mass.^{7,11} This is justified by the fact that such changes have almost no discernible effect on the bulk band structure.

If 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,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

$$E_p = \frac{3m/m_c}{2/E_g + 1/(E_g + \Delta)}. \quad (14)$$

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 data¹⁶ 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 $A_c \neq 0$.

The resulting Hamiltonian is used in an eigenvalue equation $H_8 \mathbf{F} = E \mathbf{F}$, where $\mathbf{F}(\mathbf{r})$ is a vector of envelope functions. With $A_c = 0$, the CB envelopes play a passive role, since they are completely determined by the VB envelopes of the same spin:

$$F_S = i(E - E_c)^{-1} P(k_x F_X + k_y F_Y + k_z F_Z), \quad (15)$$

which holds for both the $\uparrow$ and $\downarrow$ 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)]

$$\begin{aligned} L(E) &= L' + P^2/(E - E_c), \\ N_+(E) &= N'_+ + P^2/(E - E_c). \end{aligned} \quad (16)$$

M and N_- are unchanged. For bulk crystals the operator ordering (11) is irrelevant and one can use $N(E) = N_+(E) + N_-$. 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 $E \rightarrow E_c$. 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 $N_+ = N_- = \frac{1}{2}N$ rather than direct derivation from H_8 . This postulate leads to significant errors in the VB^{11,19} and breaks down completely near $E = E_c$. For H_6 it is better to approximate N as totally asymmetric (i.e., $N_+ \approx N$ and $N_- \approx 0$). Of course, H_8 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_{vv} , then orients the spin along an eigenvector of H_{vv} and solves H_6 . The result is a pair of identical (Kramers degenerate) cubic equations to be solved for $k^2(E)$:

$$[c_3 k^6 + c_2 k^4 + c_1 k^2 + c_0]^2 = 0, \quad (17)$$

where

$$\begin{aligned} c_3 &= LM^2 + M[(L - M)^2 - N^2]\Theta \\ &\quad + [(L - M)^3 - 3N^2(L - M) + 2N^3]\Omega, \\ c_2 &= -(\varepsilon + \frac{1}{3}\Delta)\{M(2L + M) + [(L - M)^2 - N^2]\Theta\}, \\ c_1 &= \varepsilon(\varepsilon + \frac{2}{3}\Delta)(L + 2M), \\ c_0 &= -\varepsilon^2(\varepsilon + \Delta), \end{aligned} \quad (18)$$

FIG. 1. Complex band structure of bulk $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ with $A_c = 0$ (solid lines) and $A_c \neq 0$ (dotted lines).

in which $\varepsilon = E - E_v$, $\Theta = (k_x^2 k_y^2 + k_y^2 k_z^2 + k_z^2 k_x^2)/k^4$, 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 $\varepsilon = A_H k^2$ and a quadratic equation for the split-off and light-hole/conduction bands:

$$\begin{aligned} (A_H A_L) k^4 - [\varepsilon(A_H + A_L) + \frac{1}{3}\Delta(A_H + 2A_L)] k^2 \\ + \varepsilon(\varepsilon + \Delta) = 0. \end{aligned} \quad (19)$$

For $\langle 001 \rangle$, $A_H = M$ and $A_L = L$, while for $\langle 111 \rangle$, $A_H = \frac{1}{3}(L + 2M - N)$ and $A_L = \frac{1}{3}(L + 2M + 2N)$.

These solutions are compared with the $A_c \neq 0$ 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 Δ and do not couple the spin-up and spin-down components of $\mathbf{F}$. As an example, consider a $\langle 001 \rangle$ junction. If $A_c \neq 0$, then $\mathbf{F}$ and the following functions (the same for each spin) are required to be continuous:

$$\begin{bmatrix} A_c \partial_z & 0 & 0 & 0 \\ 0 & M \partial_z & 0 & iN_- k_x \\ 0 & 0 & M \partial_z & iN_- k_y \\ P & iN'_+ k_x & iN'_+ k_y & L' \partial_z \end{bmatrix} \begin{bmatrix} F_S \\ F_X \\ F_Y \\ F_Z \end{bmatrix}. \quad (20)$$

However, if $A_c = 0$, 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

$$\begin{bmatrix} M \partial_z & 0 & iN_- k_x \\ 0 & M \partial_z & iN_- k_y \\ iN_+ k_x & iN_+ k_y & L \partial_z \end{bmatrix} \begin{bmatrix} F_X \\ F_Y \\ F_Z \end{bmatrix}, \quad (21)$$

FIG. 2. Band structure of a [001] $(\text{In}_{0.53}\text{Ga}_{0.47}\text{As})_{10}(\text{InP})_{10}$ superlattice with $A_c=0$ (solid lines); $A_c \neq 0$ and $P=\text{const}$ (dashed lines); and $A_c \neq 0$ and $P \neq \text{const}$ (dotted lines).

which is just what one obtains by integrating H_6 . Note, however, that the asymmetry (7) of H_{cv} is crucial. If H_{cv} contains any term involving $k_z P$, then the last term in the first row of (20) is not zero; it is instead proportional to P . When $A_c=0$ this forces F_Z to be discontinuous, which is not consistent with the other boundary conditions. It is for this reason that H_{cv} must have the asymmetric form (7).

The band structure of a superlattice calculated using these boundary conditions is plotted in Fig. 2. The envelope-

function equations were solved in a plane-wave basis with $|k| \leq 2.14 \text{ \AA}^{-1}$. The results for $A_c=0$ and $A_c \neq 0$ agree quite well, except where spurious bands interfere with the $A_c \neq 0$ solutions. For $A_c \neq 0$ 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 $A_c \neq 0$ 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 $A_c=0$ and $A_c \neq 0$ are due to spurious bands, since some slight differences remain even when the plane-wave basis is cut off at $|k|=0.27 \text{ \AA}^{-1}$. However, these differences are no more than those between the two $A_c \neq 0$ models ($P=\text{const}$ and $P \neq \text{const}$). The available evidence thus clearly favors the $A_c=0$ model.

The main limitation of the $A_c=0$ 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 $A_c \neq 0$ 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.

*Electronic address: phbaf@usthk.ust.hk

¹E. O. Kane, in *Handbook on Semiconductors*, edited by W. Paul (North-Holland, Amsterdam, 1982), Vol. 1, p. 193.

²G. Bastard, *Wave Mechanics Applied to Semiconductor Heterostructures* (Wiley, New York, 1988).

³S. R. White and L. J. Sham, *Phys. Rev. Lett.* **47**, 879 (1981).

⁴D. L. Smith and C. Mailhot, *Phys. Rev. B* **33**, 8345 (1986).

⁵F. Szmulowicz, *Phys. Rev. B* **54**, 11 539 (1996).

⁶M. F. H. Schuurmans and G. W. 't Hooft, *Phys. Rev. B* **31**, 8041 (1985).

⁷R. Eppenga, M. F. H. Schuurmans, and S. Colak, *Phys. Rev. B* **36**, 1554 (1987).

⁸W. Trzeciakowski, *Phys. Rev. B* **38**, 12 493 (1988).

⁹R. Winkler and U. Rössler, *Phys. Rev. B* **48**, 8918 (1993).

¹⁰C. Aversa and J. E. Sipe, *Phys. Rev. B* **49**, 14 542 (1994).

¹¹A. T. Meney, B. Gonul, and E. P. O'Reilly, *Phys. Rev. B* **50**, 10 893 (1994).

¹²M. J. Godfrey and A. M. Malik, *Phys. Rev. B* **53**, 16 504 (1996).

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.

¹⁵P. Enders, A. Bärwolff, M. Woerner, and D. Suisky, *Phys. Rev. B* **51**, 16 695 (1995).

¹⁶C. Hermann and C. Weisbuch, *Phys. Rev. B* **15**, 823 (1977).

¹⁷M. G. Burt, *J. Phys.: Condens. Matter* **4**, 6651 (1992).

¹⁸B. A. Foreman, *Phys. Rev. B* **54**, 1909 (1996).

¹⁹B. A. Foreman, *Phys. Rev. B* **48**, 4964 (1993).

²⁰J. M. Luttinger, *Phys. Rev.* **102**, 1030 (1956).

²¹In Kane's notation,¹ $N'_+ = F' - G$ and $N_- = H_1 - H_2$; in the notation of Ref. 19, $\gamma_{3+} = \sigma - \delta$ and $\gamma_{3-} = \pi$.

²²In general, the spurious solutions of H_8 are oscillatory when $A_c[L' + 2(M + N' - L')\Theta] > 0$ [where Θ is given below (18)] and evanescent when this product is negative.

²³The rapidly varying spurious solutions generate an effective discontinuity even when $A_c \neq 0$. Setting $A_c=0$ merely makes this effective discontinuity explicit.

²⁴P. von Allmen, *Phys. Rev. B* **46**, 15 382 (1992).

²⁵P. Pfeffer and W. Zawadzki, *Phys. Rev. B* **53**, 12 813 (1996).