General approach to the envelope-function approximation based on a quadrature method

R. Winkler and U. Rössler
Institut für Theoretische Physik, Universität Regensburg, D-93040 Regensburg, Federal Republic of Germany
(Received 23 December 1992)

A new concept for solving the set of coupled equations of a multicomponent envelope-function problem is presented which is based on a quadrature method for the integral equations in momentum space. It is applicable to quite general $\mathbf{k}\cdot\mathbf{p}$ Hamiltonians, considers the boundary conditions at interfaces in a natural way, and avoids automatically any spurious solutions. The flexibility and potential of the method are demonstrated by a few selected examples.

I. INTRODUCTION

Because of its central importance for fundamental physics as well as for technological applications, the electronic structure of layered semiconductor structures has attracted much interest over the last two decades.[1,2] In theoretical studies methods based on the envelope-function approximation (EFA) are predominant,[3] the reason being that the EFA allows a comprehensive description of electron and hole states. It can cope with periodic or aperiodic geometries of quantum structures as well as perturbations such as magnetic field, strain or a built-in or external potential. Details of the underlying crystal potential are included in terms of bulk band structure parameters.

The simplest approach within the EFA is the effective mass approximation[3] which assumes a single isotropic parabolic band. It provides basic insight into the electronic structure of inversion layers, heterojunctions, quantum wells and superlattices, but it fails to account for the subtleties which occur in semiconductor band structure such as nonparabolicity, spin splitting or heavy-light hole coupling. All these details of bulk band structure can be described within the framework of $\mathbf{k}\cdot\mathbf{p}$ theory. Originally developed in the early days of semiconductor physics,[4,5] it has frequently been used in applications both to bulk band structure[6–9] and to electron states in layered semiconductor structures.[10–23]

In position space the $\mathbf{k}\cdot\mathbf{p}$ matrix Hamiltonian is transformed into a kinetic energy operator by replacing $\mathbf{k}\rightarrow\frac{1}{i}\nabla$, and the corresponding Schrödinger equation is actually a set of coupled differential equations for the multicomponent envelope function.[3] The kinetic energy operator contains bulk band parameters such as matrix elements of momentum, band edge energies, effective masses and g factors. It is well defined and can be regarded as parameter free once these quantities are known, e.g., by fitting to independent experimental data obtained for the bulk material.

For layered semiconductor structures we face problems connected with the inhomogeneity of the system, i.e., interfaces with evanescent waves and changing band parameters. Commonly the latter is taken into account in the EFA by considering bulk band parameters which vary discontinuously at the interfaces.[11] As a consequence matching conditions for the envelope functions and their first derivatives have to be derived from the Hermitian form of the kinetic energy operator. The problem of finding the proper Hermitian formulation has been discussed in the literature[23,24] but remains a controversial problem.[25]

Until now no general solution to the multicomponent envelope-function problem has been provided. Several methods have been suggested, each of them suited for certain problems but always requiring additional simplifications. Hence the rest of this introduction is devoted to a short review of existing approaches, outlining their underlying physical approximations, limitations of validity, and the numerical difficulties which they entail. In the past these aspects have been mixed up to some extent.

In the bulk conduction band the effect of nonparabolicity has been taken into account by explicitly considering terms up to fourth order in $\mathbf{k}$.[26,27] This yields a simple model that has proven its worth for the conduction band in large-gap semiconductors. Furthermore, it can easily be applied to the subband problem. Commonly this is done by calculating the envelope functions from a standard effective mass Hamiltonian and then nonparabolic corrections are taken into account by first order perturbation theory.[22,28] However, difficulties may arise when dealing with expectation values of higher order in $k_z=\frac{1}{i}\partial_z$. In numerical calculations we find that these quantities can be rather large. Using generalized Fang-Howard trial functions, appropriate for metal-oxide-semiconductor (MOS) and heterostructures, one can show analytically that expectation values of all higher orders in k_z diverge.[29] In these cases a perturbative treatment is not possible.

Takada *et al.*[30] studied n-inversion layers on narrow-gap semiconductors. In an unusual approach they suggested a $\mathbf{k}$-dependent unitary transformation for diagonalizing the 6×6 multiband Hamiltonian (Γ_6^c,Γ_8^v) with respect to $E(\mathbf{k})$. In this way they were able to investigate the resonant character of subband states due to Zener tunneling into the bulk valence band. Diver-

gent expectation values of k_z were avoided by assuming $\langle k_z^{2n}\rangle \approx \langle k_z^2\rangle^n$.

Frequently a simplified multiband $\mathbf{k}\cdot\mathbf{p}$ model is used which neglects remote band contributions and the free electron term.[10–13,20,31,32] Thus it is mainly justified for narrow-gap systems. The corresponding subband problem is actually a set of coupled first order differential equations. By eliminating the valence band envelope functions it can be reduced to two second order equations solely for the conduction band envelope functions (spin up and down). However, the price to be paid is a kinetic energy operator that has singularities due to zeros of the effective mass as a function of energy and position.

White and Sham[33] and Schuurmans and 't Hooft[21] examined the solution $k^2(E)$ of the secular equation for a multiband Hamiltonian at a fixed energy. They showed that, if the Hamiltonian contains both the conduction and valence band, unphysical "wing band" or "spurious" solutions may occur. These are large imaginary or real $\mathbf{k}$ vectors which in $\mathbf{k}$ space lie far beyond the range of validity of the corresponding Hamiltonian. Likewise they occur in the case of single band Hamiltonians if they contain terms of higher than second order in $\mathbf{k}$. In contrast to many physical problems that are characterized by differential equations we have to exclude these solutions *independently of boundary conditions.* Yet the problem is that, due to numerical instabilities, these unphysical branches prevail in the standard numerical integration schemes (e.g., Runge-Kutta). This makes it impossible to solve multiband Hamiltonians along this line.

In the case of hole subbands the degeneracy of the topmost valence band Γ_8^v must be taken into account by a multicomponent formalism. When we restrict ourselves to the Luttinger Hamiltonian[5] spurious solutions do not occur. But it is still a considerable task to solve the corresponding subband problem for the four-component spinor function $\underline{\Psi}(z)$ by means of a numerical integration scheme since we have to fix the energy $\mathcal{E}$ and initial values for $\underline{\Psi}(z_0)$, $\partial_z\underline{\Psi}(z)|_{z=z_0}$ in order to determine a bound state.[34] To our knowledge this ansatz has never been realized. Instead variational procedures have been used with differently chosen trial functions.[15–19]

By using the analytical solution of the bulk secular equation Andreani, Pasquarello, and Bassani[35] showed that the eigenvalue problem with the 4×4 Luttinger Hamiltonian (Γ_8^v) can be solved exactly in the flat band case. Recently Valadares[36] and Chao and Chuang[37] extended the approach to the 6×6 valence band Hamiltonian (Γ_8^v,Γ_7^v). However, besides the restriction to a piecewise constant potential these Hamiltonians do not include nonparabolicities beyond the valence band mixing. We will show in Sec. V that these effects can be substantial even for subbands in large-gap material.

In order to overcome all these shortcomings and problems that have arisen so far in solving multicomponent EFA problems we present a concept based on a quadrature method for the coupled integral equations in momentum space. Quadrature methods have proven to be an efficient scheme for solving Schrödinger type eigenvalue problems.[38,39] The advantage of this method is that it is free from numerical difficulties as discussed above. Moreover, it can solve the EFA problem for quite general multiband $\mathbf{k}\cdot\mathbf{p}$ Hamiltonians including a built-in or external potential V in the diagonal. Proper matching of the envelope functions at the interfaces results in a natural way from the Hamiltonian. Bound and resonant states can be calculated.

We demonstrate the flexibility and strengths of our approach by applying it to a few selected subband problems. The accuracy of the method is demonstrated first by calculating the subband dispersion and envelope functions for holes in a GaAs-Al_xGa_{1-x}As quantum well and comparing these results with the exact analytical results of Andreani, Pasquarello, and Bassani.[35] The second example is the self-consistent calculation of spin-split subbands in an n-inversion layer on InSb based on an 8×8 Hamiltonian including remote band contributions. Finally we calculate the dispersion of hole subbands in a narrow GaAs-AlAs quantum well from an 8×8 Hamiltonian in order to demonstrate nonparabolicity effects due to coupling with conduction and split-off bands.

II. MULTIBAND HAMILTONIANS IN MOMENTUM SPACE

The EFA originates in the $\mathbf{k}\cdot\mathbf{p}$ eigenvalue problem for the bulk band structure[40]

$$\sum_{n'}\left\{\left[\mathcal{E}_n(\mathbf{0})+\frac{\hbar^2k^2}{2m_0}-\mathcal{E}\right]\delta_{nn'}+\frac{\hbar}{m_0}\mathbf{k}\cdot\mathbf{P}_{nn'}\right\}q_{n'}(\mathbf{k})=0. \tag{1}$$

Here $\mathcal{E}_n(\mathbf{0})$ is the band edge ($\mathbf{k}=\mathbf{0}$) energy of the nth band, $\mathbf{P}_{nn'}$ are momentum matrix elements of band edge Bloch functions and $\underline{\mathbf{Q}}(\mathbf{k})=(q_{n'})(\mathbf{k})$ is the eigenvector of expansion coefficients for the Bloch functions at $\mathbf{k}$ in terms of band edge Bloch functions. The sum over n' is infinite, but block diagonalization is used to remove coupling to remote bands. This leads to well-established models like the 4×4 Luttinger Hamiltonian[5] or the 14×14 $\mathbf{k}\cdot\mathbf{p}$ model[7,27] which now contain terms of higher order in $\mathbf{k}$.

In position space the EFA Hamiltonian is obtained from the bulk $\mathbf{k}\cdot\mathbf{p}$ matrix by replacing $\mathbf{k}\rightarrow\frac{1}{i}\nabla$ and adding the built-in or external potential $V(\mathbf{r})$ on the diagonal.[3] The resulting eigenvalue problem is a set of coupled partial differential equations for the multicomponent envelope function $\underline{\Psi}$. In the case of layered semiconductor structures (without a lateral confinement) we have $V(\mathbf{r})=V(z)$, taking the z axis in the growth direction. This makes the in-plane wave vector $\mathbf{k}_\parallel$ a good quantum number and we obtain a set of coupled ordinary differential equations.

In general, the finite-dimensional eigenvalue problem reads

$$\underline{\underline{\hat{\mathbf{H}}}}(z)\,\underline{\Psi}(z)=\mathcal{E}\,\underline{\Psi}(z). \tag{2}$$

Here the $N\times N$ matrix Hamiltonian $\underline{\underline{\hat{\mathbf{H}}}}$ acts on the spinor $\underline{\Psi}$. We have

$$\underline{\underline{\hat{\mathbf{H}}}} = \underline{\underline{\hat{\mathbf{H}}}}^{(0)} + \underline{\underline{\hat{\mathbf{H}}}}^{(1)} + \underline{\underline{\hat{\mathbf{H}}}}^{(2)} + \ldots \tag{3a}$$

with matrix elements labeled according to the band edge Bloch functions

$$\begin{aligned}
\hat{h}^{(0)}_{nn'} &= c^{(0)}_{nn'} + g^{(0)}_{nn'}(z), \\
\hat{h}^{(1)}_{nn'} &= c^{(1)}_{nn'} \frac{1}{i}\partial_z + \frac{1}{2}\left[\frac{1}{i}\partial_z\, g^{(1)}_{nn'}(z) + g^{(1)}_{nn'}(z)\,\frac{1}{i}\partial_z\right], \\
\hat{h}^{(2)}_{nn'} &= c^{(2)}_{nn'}\left(\frac{1}{i}\partial_z\right)^2 + \frac{1}{i}\partial_z\, g^{(2)}_{nn'}(z)\,\frac{1}{i}\partial_z .
\end{aligned} \tag{3b}$$

$c^{(\nu)}_{nn'} = c^{(\nu)*}_{n'n}$ are complex constants and $g^{(\nu)}_{nn'}(z) = [g^{(\nu)}_{n'n}(z)]^*$ are complex functions which describe the position dependence of the band parameters. When calculating bound states we may assume without loss of generality that the $g^{(\nu)}_{nn'}(z)$ are square integrable. In quantum wells they will be proportional to rect functions.[41] In general $\underline{\underline{\hat{\mathbf{H}}}}^{(0)}$ contains terms which result from the in-plane dispersion; in addition we have band edge energies and a potential on its diagonal. $\underline{\underline{\hat{\mathbf{H}}}}^{(1)}$ represents off-diagonal terms proportional to momentum matrix elements multiplied by $k_z = \frac{1}{i}\partial_z$ and $\underline{\underline{\hat{\mathbf{H}}}}^{(2)}$ results from remote band contributions of second order in k_z. It is straightforward to include higher orders $\underline{\underline{\hat{\mathbf{H}}}}^{(\nu)}$, $\nu > 2$ which occur, e.g., in the 2×2 Hamiltonian of Ref. 26. Equation (3b) corresponds to the most commonly used Hermitian formulation of the operator $\underline{\underline{\hat{\mathbf{H}}}}$; others are given in Table I.

Now it is straightforward to change over from position to momentum space. A Fourier transform leads to the integral equation

$$\begin{aligned}
\underline{\underline{\hat{\mathbf{H}}}}(k)\,\underline{\Psi}(k) \equiv \int_{-\infty}^{\infty} dk' \Big[& \underline{\underline{\mathbf{C}}}^{(0)}\delta(k-k') + \underline{\underline{\mathbf{G}}}^{(0)}(k-k') + \underline{\underline{\mathbf{C}}}^{(1)} k'\delta(k-k') + \frac{1}{2}(k+k')\,\underline{\underline{\mathbf{G}}}^{(1)}(k-k') \\
& + \underline{\underline{\mathbf{C}}}^{(2)} k'^2\,\delta(k-k') + kk'\,\underline{\underline{\mathbf{G}}}^{(2)}(k-k')\Big]\underline{\Psi}(k') \\
= \mathcal{E}\,\underline{\Psi}(k).
\end{aligned} \tag{4}$$

We use $k \equiv k_z$ for brevity. Hermiticity is guaranteed by the property of the integral kernel $\underline{\underline{\mathbf{H}}}(k,k') = \left[{}^T\underline{\underline{\mathbf{H}}}(k',k)\right]^*$, where ${}^T\underline{\underline{\mathbf{H}}}^*$ denotes the complex conjugate transpose of $\underline{\underline{\mathbf{H}}}$. Integral kernels corresponding to other hermitian formulations of $\underline{\underline{\hat{\mathbf{H}}}}$ are given in Table I.

The appropriate boundary conditions at interfaces emerge from the (nonunique) Hermitian formulation which one chooses for the operator $\underline{\underline{\hat{\mathbf{H}}}}$. Hence, although we need not deal with these boundary conditions in momentum space, any solution $\underline{\Psi}(k)$ of Eq. (4) will satisfy them when transformed into position space. This gives a generalization of Altarelli's interface operator method[15] which at the same time considerably simplifies our calculation.

We apply a standard quadrature method to solve Eq. (4).[42] Our ansatz is similar to the Fourier grid Hamiltonian method of Marston and Balint-Kurti[38] as well as Chao and Chuang's approach[39] to the two-dimensional exciton. Some numerical keypoints are discussed in the Appendix. A more detailed discussion will be published elsewhere.

III. COMPARISON WITH ANDREANI, PASQUARELLO, AND BASSANI

A stringent test for the accuracy of our method is to compare our results with Andreani, Pasquarello, and Bassani's exact solution of the Luttinger Hamiltonian in the flat band case.[35] Numerical values for the four-component envelope function corresponding to the highest valence subband (HH1) at $k_\parallel = 0.04$ Å^{-1} in the $\langle 10\rangle$ direction in a 100 Å GaAs-$Ga_{0.79}Al_{0.21}As$ quantum well are given in Table II. In order to illustrate the rapid convergence of the quadrature method we present calculated results for 27 and 81 grid points per spinor component (see the Appendix).

In Fig. 1 we show contour plots of the anisotropic subband dispersion corresponding to Fig. 2(a) of Ref. 35. Using a quadratic Brillouin zone integration scheme[43] we calculate the density-of-states effective mass

$$\frac{m^*(E)}{m_0} = \frac{1}{\pi}\frac{\hbar^2}{2m_0}\int \delta(E - \mathcal{E}(\mathbf{k}_\parallel))\, d^2\mathbf{k}_\parallel \tag{5}$$

TABLE I. Hermitian formulations of the kinetic energy operator (cf. Ref. 24). We compare their form in position space with the corresponding integral kernels in momentum space. For brevity we neglect the constants $c^{(\nu)}_{nn'}$. The functions $\tilde{g}^{(\nu)}_{nn'}(z)$ are defined by $[\tilde{g}^{(\nu)}_{nn'}(z)]^2 = g^{(\nu)}_{nn'}(z)$. In the table the index μ is used instead of ν to indicate that in part the labeling differs from the order of $\hat{h}^{(\nu)}_{nn'}$, cf. (ii-μ). Note that (i-1) and (iii-1) are equivalent.

Position space	Momentum space	
$\frac{1}{2}\left[(\frac{1}{i}\partial_z)^\mu\, g^{(\mu)}_{nn'}(z) + g^{(\mu)}_{nn'}(z)\,(\frac{1}{i}\partial_z)^\mu\right]$	$\frac{1}{2}(k^\mu + k'^\mu)\, g^{(\mu)}_{nn'}(k-k')$	(i-μ)
$(\frac{1}{i}\partial_z)^\mu\, g^{(2\mu)}_{nn'}(z)\,(\frac{1}{i}\partial_z)^\mu$	$(kk')^\mu\, g^{(2\mu)}_{nn'}(k-k')$	(ii-μ)
$\tilde{g}^{(\mu)}_{nn'}(z)\,(\frac{1}{i}\partial_z)^\mu\,\tilde{g}^{(\mu)}_{nn'}(z)$	$\int dk''\, k''^\mu\,\tilde{g}^{(\mu)}_{nn'}(k-k'')\,\tilde{g}^{(\mu)}_{nn'}(k''-k')$	(iii-μ)

TABLE II. Envelope function of the topmost hole subband (HH1) at $k_{\|} = 0.04$ Å^{-1} in the $\langle 10 \rangle$ direction in a 100 Å GaAs-$Ga_{0.79}Al_{0.21}As$ quantum well. We compare the exact envelope function of Andreani, Pasquarello, and Bassani (Ref. 35) with our numerical results based on the quadrature method obtained for 81 and 27 grid points per spinor component. The origin of the z axis is at the center of the quantum well.

z (Å)	ψ_1	ψ_2	ψ_3	ψ_4
		Exact envelope function (Ref. 35)		
0	7.481×10^{-2}	0.0	-1.016×10^{-1}	0.0
25	5.085×10^{-2}	1.634×10^{-2}	-7.229×10^{-2}	4.745×10^{-2}
50	7.697×10^{-3}	1.405×10^{-2}	-2.005×10^{-2}	3.077×10^{-2}
75	-1.533×10^{-3}	3.614×10^{-3}	-1.948×10^{-3}	3.797×10^{-3}
100	-6.114×10^{-4}	6.768×10^{-4}	-6.837×10^{-5}	3.306×10^{-4}
		Quadrature method: 81 grid points		
0	7.480×10^{-2}	9.296×10^{-4}	-1.016×10^{-1}	6.842×10^{-4}
25	5.084×10^{-2}	1.633×10^{-2}	-7.227×10^{-2}	4.740×10^{-2}
50	7.680×10^{-3}	1.400×10^{-2}	-2.004×10^{-2}	3.066×10^{-2}
75	-1.530×10^{-3}	3.619×10^{-3}	-1.955×10^{-3}	3.809×10^{-3}
100	-6.116×10^{-4}	6.772×10^{-4}	-6.850×10^{-5}	3.317×10^{-4}
		Quadrature method: 27 grid points		
0	7.503×10^{-2}	1.513×10^{-3}	-1.017×10^{-1}	9.376×10^{-4}
25	5.098×10^{-2}	1.636×10^{-2}	-7.228×10^{-2}	4.728×10^{-2}
50	7.895×10^{-3}	1.392×10^{-2}	-1.996×10^{-2}	3.028×10^{-2}
75	-1.444×10^{-3}	3.596×10^{-3}	-2.001×10^{-3}	3.851×10^{-3}
100	-5.928×10^{-4}	6.521×10^{-4}	-7.355×10^{-5}	3.119×10^{-4}

which is shown in Fig. 2. By comparing Fig. 2 with Fig. 2(a) in Ref. 35 it can be seen clearly that extrema in the anisotropic dispersion $\mathcal{E}(\mathbf{k}_{\|})$ correspond to Van Hove singularities in the density of states. The particular advantage of the cited $\mathbf{k}$-space integration scheme[43] is that it correctly evaluates these integrable singularities. Thus it is an important step in our self-consistent calculations discussed in Sec. IV.

IV. INVERSION ASYMMETRY INDUCED SPIN SPLITTING

Spin degeneracy of electron or hole states is a combined effect of inversion symmetry in space and time of the system under consideration. It can be removed by applying an external magnetic field which breaks time reversal symmetry.[44] A zero-field spin splitting is possible for a finite wave vector $\mathbf{k}_{\|}$ if the system lacks spatial inversion symmetry. In layered semiconductor structures this can be a consequence either of the underlying crystal (e.g., a zinc blende structure) or of the geometry of the device [e.g., inversion asymmetry of the confinement potential $V(z)$]. As far as the dispersion of electron subbands is concerned the former is known to result in a correction $\propto \alpha_1 k_{\|} + \alpha_2 k_{\|}^2 + \alpha_3 k_{\|}^3$ (Ref. 22) while the latter in lowest order will lead to a spin-orbit term[45,46]

$$H_{\mathrm{SO}} = a_{46}\, \boldsymbol{\sigma} \cdot (\mathbf{k} \times \mathbf{E}). \tag{6}$$

Here we want to focus on the so-called Rashba term (6) which is the dominant contribution to spin splitting in MOS inversion layers on narrow-gap semiconductors,[47,48] but also, for example, in asymmetric quantum wells on GaSb-InAs-GaSb.[49] Originally H_{SO} was introduced to account for the inversion asymmetry of bulk semiconductors with a wurtzite-type crystal.[45] In these systems $\mathbf{E}$ is a unit vector along the symmetry axis. Later H_{SO} was applied to inversion asymmetric semiconductor microstructures.[50,51] Here $\mathbf{E} = (0, 0, E_z)$ is an effective electric field that results from the built-in or external potential V as well as from the position dependent valence band edge.[52] The material-specific weighting factor a_{46} can be determined from the band parameters of the bulk material.[50]

In first order perturbation theory Eq. (6) results in a spin splitting strictly linear in $k_{\|}$

$$\mathcal{E}_{\mathrm{SO}}(\mathbf{k}_{\|}) = \pm \langle a_{46} E_z \rangle k_{\|}. \tag{7}$$

Using this simple formula several groups determined the prefactor $\langle a_{46} E_z \rangle$ by analyzing Shubnikov–de Haas oscillations of the magnetoresistance.[49,53] However, there are two difficulties with respect to Eq. (7).

First we find that the electric field E_z and hence also $\langle a_{46} E_z \rangle$ depends fairly sensitively on unknown properties of the oxide layer (effective mass m^*_{ox} and valence band discontinuity at the interface). This is a surprising phenomenon as the electron wave function penetrates only weakly into the barrier and thus we would have expected the band parameters used for the oxide layer to be of minor importance.

Secondly we must keep in mind that in an invariant expansion of the conduction band Hamiltonian the Rashba term (6) is merely the lowest order contribution to a spin splitting induced by the inversion asymmetry of

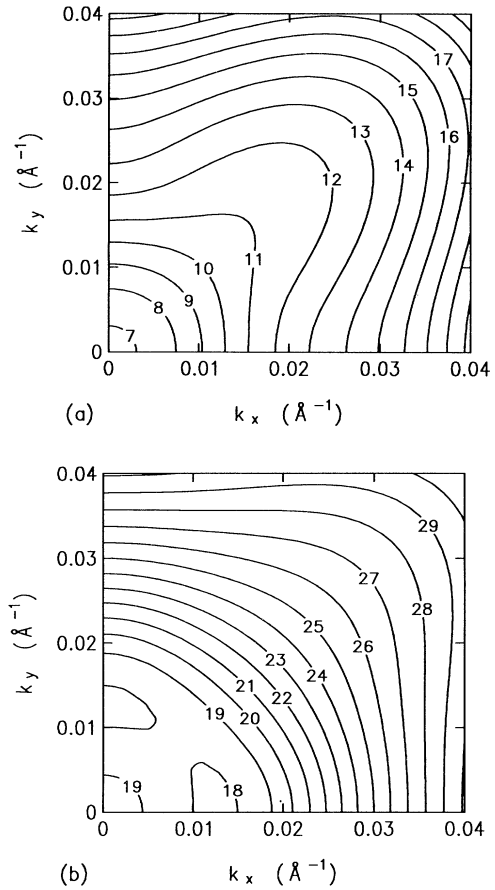

FIG. 1. Contour plots of the anisotropic hole dispersion in a 100 Å GaAs-$Al_xGa_{1-x}As$ quantum well: (a) HH1 and (b) LH1. Numbers indicate the corresponding energies in meV measured from the valence band edge. See for comparison Fig. 2(a) in Ref. 35.

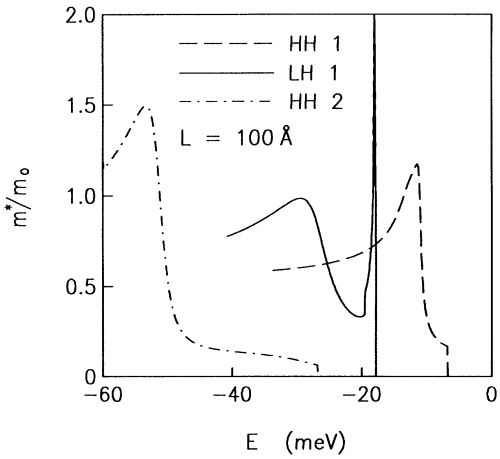

FIG. 2. Density-of-states effective mass of holes corresponding to the energy dispersion shown in Fig. 1 (this work) and Fig. 2(a) in Ref. 35.

$V(z)$. Several calculations based on simplified multiband Hamiltonians[10,12,31,32] revealed a spin splitting which is linear for small $k_\parallel$. But in contrast to Eq. (6) it appeared to be almost constant for larger values of $k_\parallel$. Until now no quantitative comparison of both approaches has been carried out. Hence it is interesting to compare the spin splitting owing to Eq. (6) [or (7)] with the results obtained from a realistic multiband Hamiltonian that fully accounts for inversion asymmetry induced spin splitting.

As an example we show in Fig. 3(a) the self-consistently calculated subband dispersion $\mathcal{E}_{i\sigma}(k_\parallel)$ for an MOS n-inversion layer on InSb with $N_S = 1.06 \times 10^{12}$ cm^{-2} and $|N_A - N_D| = 1.40 \times 10^{13}$ cm^{-3} obtained with an 8×8 Hamiltonian (Γ_6^c, Γ_8^v, and Γ_7^v) including remote band contributions.[6] Our calculation, based on the quadrature method, allows for a penetration of the envelope function into the oxide layer. The system was first investigated by Marques.[12,13]

For comparison we present in Fig. 3(b) the corresponding results obtained from a 2×2 conduction band Hamiltonian including higher order terms in $\mathbf{k}$ by perturbation theory.[22,28] The differences to Fig. 3(a) are evident: The subband edges and the Fermi level are shifted and the dispersion curves bend downwards for larger $k_\parallel$. The latter is due to the insufficiency of k^4 terms to describe nonparabolicity in narrow-gap semiconductors.

The spin splitting $|\mathcal{E}_{i\uparrow}(k_\parallel) - \mathcal{E}_{i\downarrow}(k_\parallel)|$ is shown in Fig. 3(c). For the more complete multiband Hamiltonian it is linear for small $k_\parallel$ but almost constant for larger $k_\parallel$, in good agreement with Refs. 10, 12, 31, and 32. In addition, it is insensitive to the values of the band parameters used for the oxide layer as long as they lie within a reasonable range. However, if we perform a single band calculation based either on the quadrature method and Eq. (6) or on a perturbative treatment according to Refs. 22, 28 and Eq. (7) [Figs. 3(b) and 3(c)], we obtain a spin splitting that is linear all over the range $[\mathcal{E}_0, \mathcal{E}_F]$ and that depends strongly on the band parameters used for the oxide layer.

Both aspects can be understood qualitatively by looking at a simplified 6×6 Hamiltonian (Γ_6^c, Γ_8^v) that neglects the spin split-off band Γ_7^v, remote band contributions, and the free electron term. After eliminating the valence band envelope functions we obtain a spin dependent term in the conduction band Hamiltonian[13,31]

$$H_{\rm SO}(z, k_\parallel) = \pm\frac{1}{2}\frac{\partial_z(E_v + V)}{E_v + V - \mathcal{E}}k_\parallel. \tag{8}$$

At first sight Eqs. (7) and (8) have a striking similarity. However, we may *not* simplify Eq. (8) by assuming that the denominator $D \equiv E_v + V - \mathcal{E}$ is approximately constant. Due to nonparabolicity the dispersion $\mathcal{E} = \mathcal{E}(k_\parallel)$ is almost linear for larger $k_\parallel$ so that D approximately cancels $k_\parallel$. In addition, we may assume that near the interface $z \simeq 0$ we have $E_v + V \approx V_0\,\theta(-z)$ and $|V_0| \gg \mathcal{E}$. Hence

$$H_{\rm SO}(z \simeq 0) \approx \mp\frac{1}{2}\frac{\delta(z)k_\parallel}{\theta(-z) - \mathcal{E}/V_0} \approx \mp\frac{1}{4}\delta(z)k_\parallel \tag{9}$$

independent of V_0.

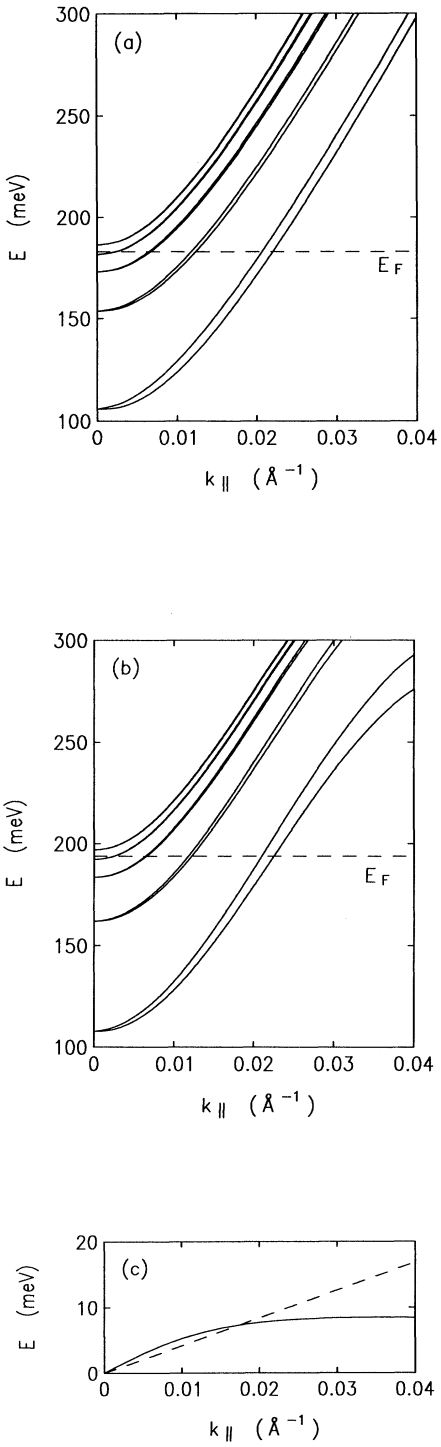

FIG. 3. Self-consistently calculated subband dispersion $\mathcal{E}_{i\sigma}(k_{\|})$ for an MOS n-inversion layer on InSb with $N_S = 1.06 \times 10^{12}$ cm^{-2} and $|N_A - N_D| = 1.40 \times 10^{13}$ cm^{-3}; (a) obtained with an 8×8 Hamiltonian (Γ_6^c, Γ_8^v, and Γ_7^v) by means of the quadrature method and (b) derived from a 2×2 Hamiltonian (Γ_6^c) that accounts for nonparabolicity and spin splitting in a perturbative way; cf. Refs. 22 and 28. (c) shows the spin splitting $|\mathcal{E}_{i\uparrow}(k_{\|}) - \mathcal{E}_{i\downarrow}(k_{\|})|$ of the lowest subband extracted from Fig. 3(a) (solid line) and Fig. 3(b) (dashed line).

The spin components $\mathcal{E}_{i\downarrow}(\mathbf{k}_{\|})$ with the lower energy have their minima shifted slightly away from $\mathbf{k}_{\|} = 0$ so that we obtain a "ring of minima" in the $\mathbf{k}_{\|}$ plane.[45] To elucidate this point we have magnified the dispersion of the lowest subband in the vicinity of $k_{\|} = 0$; see the inset of Fig. 4. Although the scale of the enlargement gives the impression that this effect is insignificant, it results in a marked Van Hove singularity in the density-of-states effective mass shown in Fig. 4. Van Hove singularities are the more pronounced the lower the dimensionality of the system under investigation. Therefore in bulk material the ring of minima is almost invisible in the density-of-states whereas in two-dimensional (2D) systems the singularities might be detectable in certain experiments.

V. HOLE SUBBANDS IN GaAs-AlAs QUANTUM WELLS

A detailed knowledge of the single particle states in layered semiconductor structures is a basic prerequisite for describing more advanced topics like optical transitions and excitons. Electron-hole pairs forming an exciton are the lowest lying excited states in semiconductor quantum wells. Since the pioneering work of Dingle, Wiegmann, and Henry[54] they have been observed in many optical experiments.

In quantum wells the confinement in growth direction removes the fourfold degeneracy of the topmost valence band Γ_8^v so that in absorption and photoluminescence excitation spectra we obtain a heavy hole–light hole (HH-LH) splitting of the exciton energy. Since many of these experiments are carried out on large-gap semiconductors (GaAs-Al_xGa_{1-x} and GaAs-AlAs) the splitting is usually explained within the framework of the 4×4 Luttinger Hamiltonian. By means of the quadrature method

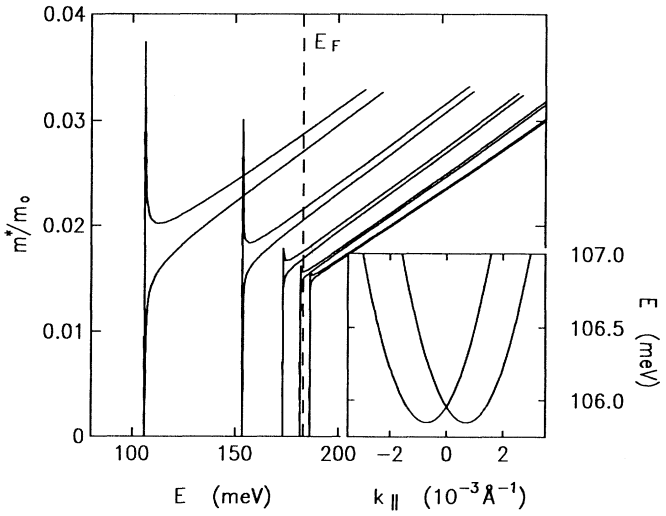

FIG. 4. Density-of-states effective mass corresponding to Fig. 3(a). Note the pronounced Van Hove singularities at the subband edges. The inset shows the dispersion of the lowest subband in the vicinity of $k_{\|} = 0$ [magnified from Fig. 3(a)].

we can easily compare the hole dispersion curves and in particular the HH-LH splitting derived from the Luttinger Hamiltonian with those obtained from more complete multiband Hamiltonians. It is well known that if we consider explicitly the off-diagonal $\mathbf{k}\cdot\mathbf{p}$ coupling between conduction and valence bands we obtain an energy shift of the subband edges $\mathcal{E}_{i\sigma}(\mathbf{k}_{\|}=0)$ and nonparabolic dispersion curves $\mathcal{E}_{i\sigma}(\mathbf{k}_{\|})$.[10,12] These effects are the more pronounced the lower the fundamental gap E_g and the narrower and deeper the quantum well. We note that the latter two conditions yield envelope functions that are widely spread in k_z space.

The system we consider is a 42 Å GaAs-AlAs quantum well. Recently Hayden *et al.*[55] used resonant tunneling spectroscopy in high magnetic fields to experimentally probe its hole dispersion curves. When they plotted the voltage position of the peaks in the tunneling current as a function of in-plane magnetic field, they observed a striking similarity to the dispersion curves calculated with the Luttinger Hamiltonian. A critical discussion of their interpretation of the experimental data will be published elsewhere.

In Fig. 5(a) we compare the hole dispersion curves derived by means of the Luttinger Hamiltonian (Γ_8^v) with those from a 6×6 Hamiltonian that explicitly contains Γ_8^v and the lowest conduction band Γ_6^c (for clarity the axial approximation is used; we found that in the case considered here anisotropy is negligible). We notice that the LH states are shifted by an amount that is of the same order of magnitude as the binding energy of a quantum well exciton. This holds even for the topmost LH state where nonparabolicity is expected to be least important. However, at $k_{\|}=0$ there is no $\mathbf{k}\cdot\mathbf{p}$ coupling between HH states and Γ_6^c so that HH subbands remain unaffected.

The spin-orbit splitting in bulk GaAs is $\Delta = 341$ meV, i.e., the bulk spin split-off (SO) valence band Γ_7^v lies well within the energy range depicted in Fig. 5(a). The extension of our envelope-function Hamiltonian with respect to Γ_7^v thus has a much stronger effect than Γ_6^c. Our results are presented in Fig. 5(b). HH states are still fully decoupled at $k_{\|}=0$ so that the subband edges $\mathcal{E}_i^{\mathrm{HH}}(k_{\|}=0)$ agree with the values derived from the Luttinger Hamiltonian (Γ_8^v). LH states, on the other hand, are shifted considerably and new subbands appear. In particular, the order of LH2 and HH3 is reversed with the consequence of a change in curvature of these subbands. Again the explicit consideration of the Γ_6^c band modifies the results as already seen in Fig. 5(a). In addition, it may change also the dominant $k_{\|}=0$ spinor component, e.g., for LH3 (SO2) in Fig. 5(b).

We conclude that the explicit consideration of the off-diagonal $\mathbf{k}\cdot\mathbf{p}$ coupling between conduction and valence bands can significantly change the subband states in quantum structures even of large-gap semiconductors. In particular, we may not restrict ourselves to the 6×6 Hamiltonian (Γ_8^v,Γ_7^v) as done recently by Valadares.[36] It can be seen from Fig. 5 that the influence of Γ_7^v and Γ_6^c on the topmost LH subband in part compensate each other so that an accurate calculation of HH-LH splitting must include both Γ_7^v and Γ_6^c. Detailed consequences for exciton states will be considered in the future.

VI. CONCLUSIONS

We have presented a new approach to the solution of multicomponent envelope-function problems which is based on a quadrature method for the integral equations in momentum space. The important advantage is that it can be applied to arbitrary $N\times N$ Hamiltonians including self-consistent potentials of quite general profile whereas in previous work these Hamiltonians had to be simplified in order to obtain a numerically tractable model. Thus we can clearly appraise the accuracy and range of validity of the simplifications used in previous publications. We have demonstrated the power of our method in applications to such different problems as (i) a quantitative comparison with the exact analytical results for a rect-

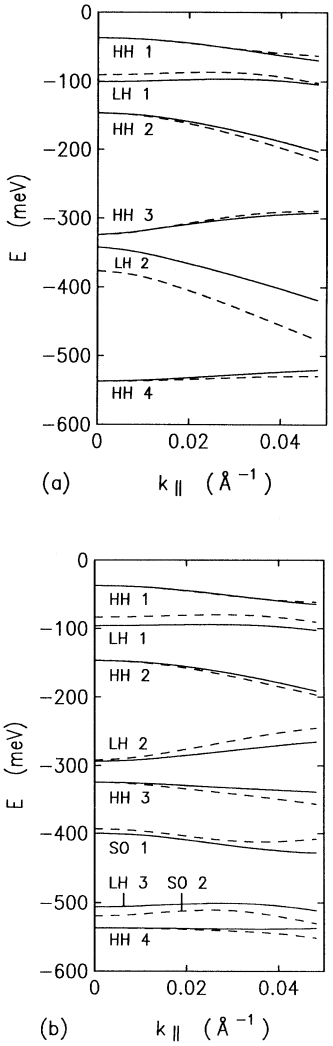

FIG. 5. Hole subband dispersion of a 42 Å GaAs-AlAs quantum well calculated by means of different multiband Hamiltonians: (a) Γ_8^v (Luttinger Hamiltonian, dashed lines), Γ_8^v and Γ_6^c (solid lines); (b) Γ_8^v and Γ_7^v (dashed lines), Γ_8^v, Γ_7^v, and Γ_6^c (solid lines). The labeling of dispersion curves corresponds to the dominant spinor component at $k_{\|}=0$.

angular quantum well, (ii) self-consistent calculation of spin-split electron subbands in an MOS inversion layer on InSb, and (iii) hole subbands in a narrow GaAs-AlAs quantum well.

Finally we want to remark that our ansatz can be readily transferred to other Schrödinger type eigenvalue problems. In particular, it is straightforward to include a quantizing magnetic field perpendicular to the interfaces.

ACKNOWLEDGMENTS

The authors want to thank F. Bolton for critically reading the manuscript. Financial support by the Deutsche Forschungsgemeinschaft is gratefully acknowledged. Numerical computations have been performed on the Cray-YMP at the Leibniz-Rechenzentrum München.

APPENDIX: NUMERICAL SOLUTION OF EQ. (4)

We apply a standard quadrature method to solve Eq. (4).[42] Quadrature methods are based on the idea that we seek to replace the calculation of an integral $I = \int_a^b \mathrm{d}x\, f(x)$ by the calculation of a finite sum of the form

$$I_m = \sum_{j=1}^{m} w_j^{(m)} f(x_j^{(m)}) \tag{A1}$$

which approximates the integral. The quantities $w_j^{(m)}$ are called the weights of the quadrature rule (A1). A sequence (I_m) of quadrature rules represents a quadrature method. Riemann sums are a well-known example for a quadrature method. They result from a sequence of partitions $P_m = \{p_j^{(m)} : a = p_0^{(m)} \le p_1^{(m)} \le \cdots \le p_{m-1}^{(m)} \le p_m^{(m)} = b\}$ with $w_j^{(m)} = p_j^{(m)} - p_{j-1}^{(m)}$ and the grid points $x_j^{(m)}$ are chosen such that $p_{j-1}^{(m)} \le x_j^{(m)} \le p_j^{(m)}$. By varying the number of grid points m one can easily check the accuracy of the numerical results.

The idea of Eq. (A1) can be applied to eigenvalue integral equations[42] of the form

$$\int_a^b \mathrm{d}x'\, k(x,x')\,\psi(x') = \lambda\,\psi(x). \tag{A2}$$

We make x and x' discrete so that the integral kernel $k(x,x')$ becomes a finite matrix. In this way we obtain a matrix eigenvalue problem (Nyström's method[42])

$$\sum_{j=1}^{m} w_j^{(m)}\, k(x_i^{(m)}, x_j^{(m)})\,\psi(x_j^{(m)}) = \lambda\,\psi(x_i^{(m)}) \tag{A3}$$

which can be solved numerically. It is straightforward to extend these formulas so that we can treat a set of coupled integral equations.[42]

The procedure outlined above is well-established for the case of integral equations like (A2). But the integral equations (4) differ from the "usual" Fredholm equations (A2) for two reasons. Firstly the domain of integration is infinite and secondly our integral kernel is an unbounded operator. Thus it is not possible to prove convergence of the quadrature rules in the usual way.[56,57] To the best of our knowledge little research has been published on this class of integral equations.

The range of validity of the commonly used $\mathbf{k}\cdot\mathbf{p}$ Hamiltonians is restricted to $|k| \ll 2\pi/a$ (a is the lattice constant). Accordingly we are interested in those solutions of Eq. (4) that fall off rapidly outside a small interval centered about $k = 0$. Thus the use of the finite section approximation[56] is justified in which the limits of integration are replaced by finite numbers. In the case of Wiener-Hopf and related integral equations it has been shown that the approximate solutions converge uniformly on every finite interval.[56] On the finite domain of integration the integral operator is compact and the quadrature method yields eigenfunctions in the space $\mathcal{L}^2$ when (A3) is diagonalized. Wing band and spurious solutions correspond to large $|k|$ values far outside the domain of integration that we use in the finite section approximation. Therefore these unphysical solutions do not occur in our approach.[58]

Even in those cases where unphysical solutions are impossible (e.g., when we restrict ourselves to the 4×4 Luttinger Hamiltonian[35]) we find that the finite section approximation is more suitable than mapping the infinite domain of integration onto a finite interval by a transformation of variables because in many applications we want to change between position and momentum representation by means of a discrete Fourier transform. Therefore the quadrature rules have to provide the eigenfunctions on a set of equidistant grid points.[59]

The situation is somewhat difficult if in $\underline{\underline{\hat{\mathbf{H}}}}$ we include higher order terms $\mu > 1$ from Table I. For abrupt interfaces, i.e., in case of quantum wells $g_{nn'}^{(\nu)}(k-k') \propto \sin[l(k-k')/2]/(k-k')$ (here l is the width of the quantum well), only the kernels $\mu = 1$ remain finite when the integrand tends to infinity. The divergence for $\mu \ge 2$ is intimately related to the argument used by Morrow and Brownstein[24] to rule out the Hermitian form (i-2) of Table I. One can show that at an abrupt interface an envelope function corresponding to (i-2) must have a δ-function-like integrable singularity. As of yet we do not know whether from a mathematical point of view the eigenvalue problem Eq. (4) is well defined for abrupt interfaces and arbitrary kernels $\mu > 1$.

A numerical check of convergence is possible if we consider analytically solvable models such as the Luttinger Hamiltonian in the flat band case (see Sec. III). We find that we can easily obtain numerical results of arbitrary accuracy if we sufficiently increase the number of grid points m which define the quadrature rules (see Table II). In so doing the energy eigenvalues converge even faster then the eigenfunctions.

The Dirac notation provides a slightly different view on our method, thus emphasizing its generality. We can expand the eigenvalue problem $\hat{H}\,|\psi\rangle = E\,|\psi\rangle$ either in terms of the complete set of position eigenstates $\{|\mathbf{r}\rangle\}$ (which in numerical calculations is the more usual approach)

$$\int d\mathbf{r}' \langle \mathbf{r}|\hat{H}|\mathbf{r}'\rangle\langle \mathbf{r}'|\psi\rangle = E\langle \mathbf{r}|\psi\rangle \tag{A4a}$$

or in terms of the likewise complete set of momentum eigenstates $\{|\mathbf{k}\rangle\}$

$$\int d\mathbf{k}' \langle \mathbf{k}|\hat{H}|\mathbf{k}'\rangle\langle \mathbf{k}'|\psi\rangle = E\langle \mathbf{k}|\psi\rangle. \tag{A4b}$$

We may choose either representation depending on which is more convenient for the evaluation of matrix elements of $\hat{H}$. In both cases we will approximate the eigenvalue problem by considering only a finite number m of matrix elements. Thus both approaches converge equally well if m is increased.

[1] T. Ando, A. B. Fowler, and F. Stern, Rev. Mod. Phys. **54**, 437 (1982).

[2] G. Bastard, J. A. Brum, and R. Ferreira, Solid State Phys. **44**, 229 (1991).

[3] G. Bastard, *Wave Mechanics Applied to Semiconductor Heterostructures* (Les Editions de Physique, Les Ulis, Cedex, 1988).

[4] E. O. Kane, J. Phys. Chem. Solids **1**, 82 (1956); **1**, 249 (1957).

[5] J. M. Luttinger and W. Kohn, Phys. Rev. **97**, 869 (1955); J. M. Luttinger, *ibid.* **102**, 1030 (1956).

[6] H.-R. Trebin, U. Rössler, and R. Ranvaud, Phys. Rev. B **20**, 686 (1979).

[7] U. Rössler, Solid State Commun. **49**, 943 (1984).

[8] M. Cardona, N. E. Christensen, and G. Fasol, Phys. Rev. Lett. **56**, 2831 (1986).

[9] M. Cardona, N. E. Christensen, and G. Fasol, Phys. Rev. B **38**, 1806 (1988).

[10] F. J. Ohkawa and Y. Uemura, J. Phys. Soc. Jpn. **37**, 1325 (1974).

[11] G. Bastard, Phys. Rev. B **24**, 5693 (1981); **25**, 7584 (1982).

[12] G. E. Marques and L. J. Sham, Surf. Sci. **113**, 131 (1982).

[13] G. E. Marques, Ph.D. thesis, University of California, 1982.

[14] W. Zawadzki, J. Phys. C **16**, 229 (1983).

[15] M. Altarelli, in *Application of High Magnetic Fields in Semiconductor Physics*, edited by G. Landwehr, Lecture Notes in Physics Vol. 177 (Springer, Berlin, 1983), p. 174.

[16] U. Ekenberg and M. Altarelli, Phys. Rev. B **30**, 3569 (1984); **32**, 3712 (1985).

[17] M. Baumgartner, G. Abstreiter, and E. Bangert, J. Phys. C **17**, 1617 (1984).

[18] D. A. Broido and L. J. Sham, Phys. Rev. B **31**, 888 (1985).

[19] T. Ando, J. Phys. Soc. Jpn. **54**, 1528 (1985).

[20] R. Lassnig, Phys. Rev. B **31**, 8076 (1985).

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

[22] F. Malcher, G. Lommer, and U. Rössler, Superlatt. Microstruct. **2**, 267 (1986); **2**, 273 (1986).

[23] M. G. Burt, J. Phys. Condens. Matter **4**, 6651 (1992).

[24] R. A. Morrow and K. R. Brownstein, Phys. Rev. B **30**, 678 (1984).

[25] B. Laikhtman, Phys. Rev. B **46**, 4769 (1992).

[26] M. Braun and U. Rössler, J. Phys. C **18** 3365 (1985).

[27] H. Mayer and U. Rössler, Phys. Rev. B **44**, 9048 (1991).

[28] R. Winkler, U. Kunze, and U. Rössler, Surf. Sci. **263**, 222 (1992).

[29] Unpublished results based on Eq. (12) in Ref. 28.

[30] Y. Takada, K. Arai, N. Uchimura, and Y. Uemura, J. Phys. Soc. Jpn. **49**, 1851 (1980); Y. Takada, K. Arai, and Y. Uemura, in *Physics of Narrow Gap Semiconductors*, edited by E. Gornik, H. Heinrich, and L. Palmetshofer, Lecture Notes in Physics Vol. 152 (Springer, Berlin, 1982), p. 101.

[31] I. Nachev, Semicond. Sci. Technol. **3**, 29 (1988).

[32] P. Sobkowicz, Semicond. Sci. Technol. **5**, 183 (1990).

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

[34] *A posteriori* one of the initial values appears to be unnecessary due to the normalization of the envelope functions.

[35] L. C. Andreani, A. Pasquarello, and F. Bassani, Phys. Rev. B **36**, 5887 (1987).

[36] E. C. Valadares, Phys. Rev. B **46**, 3935 (1992).

[37] C. Y.-P. Chao and S. L. Chuang, Phys. Rev. B **46**, 4110 (1992).

[38] C. C. Marston and G. G. Balint-Kurti, J. Chem. Phys. **91**, 3571 (1989).

[39] C. Y.-P. Chao and S. L. Chuang, Phys. Rev. B **43**, 6350 (1991).

[40] E. O. Kane, in *Semiconductors and Semimetals I*, edited by R. K. Willardson and A. C. Beer (Academic, New York, 1966), p. 75.

[41] The rect function is defined by

$$\mathrm{rect}\,(z) = \begin{cases} 1, & |z| \le 1 \\ 0, & |z| > 1. \end{cases}$$

[42] C. T. H. Baker, *The Numerical Treatment of Integral Equations* (Clarendon, Oxford, 1977).

[43] R. Winkler, J. Phys. Condens. Matter **5**, 2321 (1993).

[44] C. Kittel, *Quantum Theory of Solids* (Wiley, New York, 1963).

[45] E. I. Rashba, Fiz. Tverd. Tela (Leningrad) **2**, 1224 (1960) [Sov. Phys. Solid State **2**, 1109 (1960)].

[46] Y. A. Bychkov and E. I. Rashba, J. Phys. C **17**, 6039 (1984).

[47] A. Därr, J. P. Kotthaus, and T. Ando, in *Proceedings of the 13th International Conference on the Physics of Semiconductors, Rome, 1976*, edited by F. G. Fumi (North-Holland, Amsterdam, 1976), p. 774.

[48] G. Lommer, F. Malcher, and U. Rössler, Phys. Rev. Lett. **60**, 728 (1988).

[49] J. Luo, H. Munekata, F. F. Fang, and P. J. Stiles, Phys. Rev. B **41**, 7685 (1990).

[50] G. Lommer, F. Malcher, and U. Rössler, Phys. Rev. B **32**, 6965 (1985).

[51] U. Rössler, F. Malcher, and G. Lommer, in *High Magnetic Fields in Semiconductor Physics II*, edited by G. Landwehr, Solid-State Sciences Vol. 87 (Springer, Berlin, 1989), p. 376

[52] In previous work (Refs. 22, 47, and 48) it was overlooked that we have to evaluate $\partial_z[E_v(z) + V(z)]$ and *not* $\partial_z[E_c(z) + V(z)]$. This was pointed out by Lassnig (Ref. 20).

[53] R. Wollrab, R. Sizmann, F. Koch, J. Ziegler, and H. Maier, Semicond. Sci. Technol. **4**, 491 (1989).

[54] R. Dingle, W. Wiegmann, and C. H. Henry, Phys. Rev. Lett. **33**, 827 (1974),

[55] R. K. Hayden, D. K. Maude, L. Eaves, E. C. Valadares, M. Henini, F. W. Sheard, O. H. Hughes, J. C. Portal, and L. Cury, Phys. Rev. Lett. **66**, 1749 (1991); R. K. Hayden, L. Eaves, M. Henini, D. K. Maude, J. C. Portal, and G. Hill, Appl. Phys. Lett. **60**, 1474 (1992).

[56] P. M. Anselone and I. H. Sloan, J. Integral Eq. **9** (Suppl.), 3 (1985); F. R. de Hoog and I. H. Sloan, J. Austral. Math. Soc. Ser. B **28**, 415 (1987).

[57] I. H. Sloan, Math. Comp. **36**, 511 (1981); P. M. Anselone and I. H. Sloan, J. Integral Eq. and Appl. **1**, 203 (1988); G. A. Chandler and I. G. Graham, Numer. Math. **52**, 345 (1988).

[58] At first sight one might argue that the calculated eigenfunctions will be square integrable simply because we make a Fourier transform of Eq. (3) in order to obtain the eigenvalue problem (4) and a Fourier transform is restricted to the space $\mathcal{L}^2$. However, *first* we establish the integral operator (the Hamiltonian) and *afterwards* we determine its eigenfunctions.

[59] E. O. Brigham, *The Fast Fourier Transform* (Prentice-Hall, Englewood Cliffs, NJ, 1974).